THE ACCOUNTING REVIEW

C06-H02564-36-P028879

OLUME 81, No. 2 MARCH 2006

AJIV D. BANKER and LEI (TONY) CHEN
Predicting Earnings Using a Model Based on Cost Variability and Cost Stickiness

AQUELINE S. HAMMERSLEY
Pattern Identification and Industry-Specialist Auditors

ESLIE D. HODDER, PATRICK E. HOPKINS, and JAMES M. WAHLEN
Risk-Relevance of Fair-Value Income Measures for Commercial Banks

ATHRYN KADOUS, SUSAN D. KRISCHE, and LISA M. SEDOR
Using Counter-Explanation to Limit Analysts' Forecast Optimism

ANJANI KRISHNAN, MICHELLE H. YETMAN, and ROBERT J. YETMAN
Expense Misreporting in Nonprofit Organizations

HI-CHUN LIU and STEPHEN G. RYAN
Income Smoothing over the Business Cycle: Changes in Banks' Coordinated Management of Provisions for Loan Losses and Loan Charge-Offs from the Pre-1990 Bust to the 1990s Boom

TEPHEN G. RYAN, JENNIFER W. TUCKER, and PAUL A. ZAROWIN
Classification and Market Pricing of the Cash Flows and Accruals on Trading Positions

REMILA GOWRI SHANKAR and HUN-TONG TAN
Determinants of Audit Preparers' Workpaper Justifications

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2005–2006 Executive Committee of the Association are as follows:

President	Judy D. Rayburn, University of Minnesota
President-Elect	Shyam Sunder, Yale University
Vice President	James E. Hunton, Bentley College
Vice President–Education	Nancy A. Bagranoff, Old Dominion University
Vice President–Finance	Susan F. Haka, Michigan State University
Vice President–International	Alfred Wagenhofer, Universitat Graz
Vice President–Professional Relations	D. Scott Showalter, KPMG
Vice President–Publications	Morton P. Pincus, University of California, Irvine
Vice President–Research	John C. Fellingham, The Ohio State University
Vice President–Sections and Regions	Susan V. Crosson, Santa Fe Community College
Past President	Jane F. Mutchler, Georgia State University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive Accounting Education News, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $145.00	with Electronic Option $165.00
Select two Association-wide journals	U.S. $155.00	with Electronic Option $175.00
All three Association-wide journals	U.S. $165.00	with Electronic Option $185.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $25.00	with Electronic Option $45.00
Select two Association-wide journals	U.S. $35.00	with Electronic Option $55.00
All three Association-wide journals	U.S. $45.00	with Electronic Option $65.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to TARUAZ@email.arizona.edu or sent to Professor Dan Dhaliwal, Senior Editor. *The Accounting Review,* The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108.

The submission fee of $125.00 for members of the AAA or $200.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published five times a year in January, March, May, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $300; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review,* 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2006 American Accounting Association. All rights reserved. Printed in the United States of America.

36

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Dan S. Dhaliwal
Editors
John E. Core
Steven J. Huddart
William R. Kinney, Jr.
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan

P28879

VOL. 81 MARCH 2006 No. 2

RAJIV D. BANKER and LEI (TONY) CHEN
Predicting Earnings Using a Model Based on Cost Variability and Cost Stickiness .. 285
JAQUELINE S. HAMMERSLEY
Pattern Identification and Industry-Specialist Auditors 309
LESLIE D. HODDER, PATRICK E. HOPKINS, and JAMES M. WAHLEN
Risk-Relevance of Fair-Value Income Measures for Commercial Banks 337
KATHRYN KADOUS, SUSAN D. KRISCHE, and LISA M. SEDOR
Using Counter-Explanation to Limit Analysts' Forecast Optimism 377
RANJANI KRISHNAN, MICHELLE H. YETMAN, and ROBERT J. YETMAN
Expense Misreporting in Nonprofit Organizations 399
CHI-CHUN LIU and STEPHEN G. RYAN
Income Smoothing over the Business Cycle: Changes in Banks' Coordinated Management of Provisions for Loan Losses and Loan Charge-Offs from the Pre-1990 Bust to the 1990s Boom 421
STEPHEN G. RYAN, JENNIFER W. TUCKER, and PAUL A. ZAROWIN
Classification and Market Pricing of the Cash Flows and Accruals on Trading Positions 443
PREMILA GOWRI SHANKAR and HUN-TONG TAN
Determinants of Audit Preparers' Workpaper Justifications 473
Editorial Data 496
Editorial Policy and Style Information 497

NO-2,3

Editorial Staff of *The Accounting Review*
2005–2006

SENIOR EDITOR
Dan S. Dhaliwal, The University of Arizona and The University of Auckland
EDITORS
John E. Core
Steven J. Huddart
William R. Kinney, Jr.
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan
EDITORIAL ASSISTANT
Gail Emily Fey, The University of Arizona

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Judy Cothern
Jan Kovarik (subcontracted)
Marcia Craig (subcontractor)
OTHER AAA STAFF
Debora Gardner
Peggy Turczyn
Linda Watterworth
Pat Stein
Deirdre Harris (subcontracted)
Roland LaTulip
James Szelminski
Mary Cole
Barbara Brady
Myriam Chacon

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Anil Arya, The Ohio State University
Hollis (Holly) Ashbaugh-Skaife, University of Wisconsin–Madison
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, The Pennsylvania State University
Sudipta Basu, Emory University
Anne L. Beatty, The Ohio State University
Messod Daniel Beneish, Indiana University
Daniel A. Bens, The University of Arizona
Jennifer Blouin, University of Pennsylvania
Robert J. Bloomfield, Cornell University
James R. Boatsman, Arizona State University
Christine A. Botosan, University of Utah
Mark T. Bradshaw, Harvard University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
William M. Cready, The University of Texas at Dallas
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Merle M. Erickson, The University of Chicago
Leslie G. Eldenburg, The University of Arizona
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
Jennifer J. Gaver, University of Georgia
Frank B. Gigler, University of Minnesota, Twin Cities Campus
Dan Givoly, The Pennsylvania State University
Cristi A. Gleason, The University of Iowa
Jonathan C. Glover, Carnegie Mellon University
Michelle Hanlon, University of Michigan
R. Lynn Hannan, Georgia State University
Rachel M. Hayes, University of Utah
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
S. Paul Hribar, Cornell University
Christopher D. Ittner, University of Pennsylvania
Bruce W. Johnson, The University of Iowa
Karla M. Johnstone, The University of Wisconsin–Madison
Kathryn Kadous, Emory University
Sanjay Kallapur, Purdue University and the Indian School of Business
Steven J. Kachelmeier, The University of Texas at Austin
Bin Ke, The Pennsylvania State University
Elizabeth K. Keating, Harvard University
Inder K. Khurana, University of Missouri–Columbia
Ronald R. King, Washington University in St. Louis
April Klein, New York University
Lisa L. Koonce, The University of Texas at Austin
Ranjani Krishnan, Michigan State University
Linda K. Krull, The University of Texas at Austin
Christian Leuz, University of Pennsylvania
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Robert C. Lipe, University of Oklahoma
Oliver Zhen Li, University of Notre Dame
Laureen A. Maines, Indiana University
Carol A. Marquardt, New York University
Dawn A. Matsumoto, University of Washington
Edward L. Maydew, The University of North Carolina at Chapel Hill
Michael B. Mikhail, Duke University
Gregory S. Miller, Harvard University
Lillian F. Mills, The University of Arizona
Partha Mohanram, Columbia University
Donald Moser, University of Pittsburgh
James N. Myers, Texas A&M University
Venky Nagar, University of Michigan
Karen K. Nelson, Rice University
Kaye J. Newberry, University of Houston

Doron Nissim, Columbia University
Patricia C. O'Brien, University of Waterloo
Per M. Olsson, Duke University
Mark E. Peecher, University of Illinois at Urbana–Champaign
Kathy R. Petroni, Michigan State University
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, The University of Chicago
Marlene A. Plumlee, University of Utah
Gordon S. Potter, Cornell University
Grace Pownall, Emory University
Suresh Radhakrishnan, The University of Texas at Dallas
Shivaram Rajgopal, University of Washington
Scott A. Richardson, University of Pennsylvania
Darren T. Roulstone, The University of Chicago
Rachel Schwartz, University of Illinois at Urbana–Champaign
K. (Shiva) Sivaramakrishnan, University of Houston
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Phillip C. Stocken, Dartmouth College
Hun-Tong Tan, Nanyang Technology University/Singapore
Robert H. Trezevant, University of Southern California
Mark Trombley, The University of Arizona
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Richard H. Willis, Tulane University
Peter D. Wysocki, Massachusetts Institute of Technology
Hong Xie, University of Illinois at Urbana–Champaign
Xiao-Jun Zhang, University of California, Berkeley

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1960–1962	Robert K. Mautz
1963–1964	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William R. Kinney, Jr.
1990–1994	A. Rashad Abdel–khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 285–307

Predicting Earnings Using a Model Based on Cost Variability and Cost Stickiness

Rajiv D. Banker
Temple University

Lei (Tony) Chen
Georgia State University

ABSTRACT: We evaluate the descriptive validity of the cost behavior model for profit analysis using Compustat data. For this purpose, we propose an earnings forecast model decomposing earnings into components that reflect (1) variability of costs with sales revenue and (2) stickiness in costs with sales declines. We evaluate the predictive ability of our model by benchmarking its performance in forecasting one-year-ahead returns on equity against that of two other time-series models based on line item information reported in the income statement and in the statement of cash flows. Specifically, we consider a model that disaggregates earnings into operating income and non-operating income components and another that disaggregates earnings into cash flows and accruals components. While all three models are less accurate than analysts' consensus forecasts that rely on a larger information set, we find that our model provides substantial improvement in forecast accuracy over the other two models that use only the line items in the financial statements. Finally, invoking the market efficiency assumption, we find that earnings forecast errors based on our model have greater relative information content than forecast errors based on the two alternative models based on financial statement information in explaining abnormal stock returns.

Keywords: *analyst forecasts; cost variability; earnings forecasts; fixed costs; market association; predictive content; sticky costs; variable costs.*

Data Availability: *Data used in this study are available from public sources identified in the paper.*

I. INTRODUCTION

Understanding cost behavior is one of the most important aspects of profit analysis for managers. Central to the cost-volume-profit analysis discussed in most managerial accounting textbooks is the traditional model of fixed and variable costs

---

We gratefully acknowledge helpful comments and suggestions from participants in research seminars at University of Arkansas, University of California, San Diego, Michigan State University, Rice University, Temple University, and The University of Texas at Dallas, from the discussants of our paper and other participants at the 2004 Management Accounting Section and the 2004 Financial Accounting and Reporting Section Midyear Meetings, and from two anonymous reviewers. We also thank Thomson Financial I/B/E/S for providing the data on analysts' earnings forecasts.

Editor's note: This paper was accepted by Lawrence Brown, Editor.

Submitted July 2003
Accepted August 2005

(Garrison and Noreen 2002). If this model is descriptively valid, then its estimation using past data should provide a basis for forecasting future earnings. Despite the conceptual importance of incorporating cost behavior in conducting profitability analysis, few empirical studies have systematically examined the forecasting ability of models that explicitly recognize the relation between costs and sales when forecasting future profits. Our purpose is to evaluate the performance of such a cost behavior model in forecasting earnings by benchmarking the predictive performance of our model against that of alternative time-series models using earnings component classifications reported in the income statement and in the statement of cash flows.[1]

Prior research addresses the question of forecasting earnings from an external reporting perspective; that is, earnings forecast models are based on line items reported in the income statement (Lipe 1986; Fairfield et al. 1996), on accounting signals calculated from financial statements (Ou and Penman 1989), or on components of earnings such as cash flows and accruals (Sloan 1996). We propose an approach grounded in the management accounting tradition that models earnings as consisting of components proportional to sales increases and decreases, and another unrelated to sales changes. This method of earnings decomposition is motivated by a cost behavior model that recognizes sales as the key driver of profit and variable costs as varying with sales. In addition, we incorporate in our model evidence that cost changes exhibit stickiness in periods of sales decline (Noreen and Soderstrom 1997; Anderson et al. 2003). Cost stickiness refers to the phenomenon that costs decrease less with a sales decrease than they increase with an equivalent sales increase. Economic considerations such as adjustment costs of reducing capacity and ramping up capacity in case sales rebound, motivate managers not to cut back on activity resources in response to a decline in sales to the same extent that they increase resources in response to an increase in sales (Anderson et al. 2003). This asymmetric cost behavior implies asymmetric behavior of earnings with respect to sales changes. Our cost behavior based approach to forecasting earnings is consistent with recent calls to incorporate economic considerations in models characterizing earnings components (Kothari 2001, 151).

Our model based on *cost variability and cost stickiness* (CVCS) is informationally parsimonious, relying only on earnings and sales time-series data. To evaluate the predictive ability of our CVCS model based on information drawn from the financial statements, we benchmark its performance against four other forecast models, including two that rely on information on past earnings components to predict future returns on equity (ROE).[2] The first model (Fairfield et al. 1996) uses income statement line items to classify earnings into operating income and non-operating income components to generate forecasts of future ROE (hereafter, the OPINC model). The second model (Sloan 1996) predicts future ROE based on past cash flows and accruals components of earnings (hereafter, the CASHFLOW model). The interest in earnings decomposition in forecasting stems largely from the belief that earnings components have different time-series properties and incorporating such information through earnings disaggregation provides a more accurate forecast of future earnings (Fairfield et al. 1996; Kothari 2001). To provide a baseline benchmark, we also consider a simple forecast model based only on past ROE (hereafter, the ROE model). Prior research provides evidence that analysts' earnings forecasts are more accurate than time-series models due to their information advantage in collecting, processing, and aggregating multiple sources of financial and nonfinancial information for forecasting earnings (Brown et al. 1987a; O'Brien 1988). Therefore, we also evaluate the performance of models based on earnings disaggregation relative to that of analysts' consensus forecast to gauge the extent to which different methods of earnings decomposition succeed in capturing some of the analysts' information advantage.

---

[1] We describe our proposed model and other benchmark models as "time-series" models because they are based on a time-series correlation between current and past realizations of components of earnings. However, the empirical estimation of these models may be done on a cross-sectional or pooled basis, rather than on separate time-series data of individual firms to improve precision of the estimated model coefficients.

[2] We focus on forecasting ROE to be consistent with prior literature (Fairfield et al. 1996; Frankel and Lee 1998). However, results not reported in the paper indicate that the documented superior forecasting performance of our CVCS model relative to other time-series models we consider is robust to forecasting return on assets (ROA).

Our primary objective is to evaluate whether models that explicitly incorporate the relationship between cost changes and sales changes are better in capturing the behavior of the earnings time-series than do models that ignore such a relationship between income statement items.[3] Using a sample of 8,771 firms in the period 1992–2002, we provide evidence that our proposed CVCS model substantially improves forecast accuracy of one-year-ahead ROE over the ROE, OPINC, and CASHFLOW models. Specifically, we find that the reduction in median absolute forecast errors using our CVCS model relative to the ROE model is about nine times as large as the reduction in median absolute forecast errors using the OPINC or the CASHFLOW model relative to the ROE model. We also find that recognizing cost variability and cost stickiness in earnings decomposition reduces over 30 percent of the advantage in median absolute forecast errors that financial analysts' forecasts display relative to the ROE model.[4]

Invoking the market efficiency hypothesis, we examine whether our CVCS model has greater relative information content in capturing the market's earnings expectations than the other three time-series models. Relative information content is assessed based on the relative strength of the association between forecast errors in ROE from each of these models and the contemporaneous abnormal stock returns. Using a sample of 4,348 firms in the period 1992–2002, we find that our CVCS model also outperforms the other three models in terms of relative information content. The evidence suggests the information in earnings components that pertains to cost variability and cost stickiness is correlated with a part of the information set used in forming the capital market's earnings expectations.

We organize the remainder of the paper as follows. In the next section we discuss earnings forecasting using earnings components and present our contrasting method of earnings decomposition based on cost variability and cost stickiness. We describe the sample and forecast models in Section III, and we report results of forecast accuracy comparisons in Section IV. In Section V, we present tests of relative information content of our model in evaluating the association between earnings forecast errors and abnormal stock returns. Finally, in Section VI we offer concluding remarks and suggest directions for future research.

II. EARNINGS DECOMPOSITION AND FORECASTING

Earnings Forecasting Using Earnings Components

Our study is related to a branch of the earnings forecasting literature concerned with the use of earnings components in predicting future earnings (Brown 1993). Fairfield et al. (1996) and Kothari (2001) suggest that partitioning earnings into components with different levels of persistence may enable superior modeling of the earnings process and improve the accuracy of earnings forecasts. Two time-series forecast models using past earnings components to predict future earnings have attracted much attention in the prior literature.[5] One such model decomposes current earnings into cash flows and accruals components and predicts future earnings using them. Sloan (1996) provides evidence that cash flows have higher persistence than accruals for assessing one-year-ahead earnings but he does not examine the out-of-sample forecasting performance of this model because his focus is on whether the market fully incorporates the information about the differential persistence of cash flows and accruals components. The other earnings-decomposition-based forecast model uses line items in the income statement as the basis for predicting future earnings. Fairfield et al. (1996) find that decomposing earnings into operating income and non-operating income improves the accuracy of one-year-ahead forecasts of returns on equity, relative to a basic model using only aggregate earnings, although further disaggregation does not result in more accurate earnings forecasts.

---

[3] We restrict our attention to parsimonious models based only on income statement and cash flow statement data, and do not attempt to develop forecast models (analogous to analysts' forecast models) relying on a wide array of information beyond just financial statements. However, the model structure we develop to link sales changes to cost changes, and consequently to earnings changes, may be augmented to incorporate economy-, industry-, and firm-specific information affecting sales or cost time-series.

[4] The corresponding reduction in median absolute forecast errors is less than 6 percent for the OPINC and CASHFLOW models.

While these two methods of decomposition seem natural and intuitive from the standpoint of financial reporting, they do not fully exploit the information about the relationships among earnings components suggested by the underlying economics of firms' production processes. We propose an alternative approach to earnings decomposition that reflects the earnings generating process as the outcome of a more primitive sales-generating process. At the operational level, this approach is derived from a management accounting perspective that views accounting income as sales revenue net of variable and fixed expenses, with the explicit recognition of the relationship between revenue and variable expenses. To the extent that variable costs are driven by sales revenue, accounting earnings may be modeled as consisting of two components, the first corresponding to contribution margin varying with sales, and the second corresponding to fixed costs uncorrelated with sales. The focus on sales as the key profit driver allows us to model the earnings process as reflecting the underlying economics of the firms' process of converting resources (costs) into outputs (sales). By relying on such a conceptual framework that highlights the intercorrelation of earnings components, we can incorporate into our forecast model potential asymmetric cost behavior related to sales changes. This constitutes a fundamental difference between our approach and prior earnings-decomposition-based forecast models. In the next section, we discuss in detail how we model such asymmetry (i.e., cost stickiness).

A similar perspective focusing on the prominent role of sales as the driver of earnings is adopted by Dechow et al. (1998), who model earnings, cash flows, and accruals as stochastic processes conditional on a random walk sales-generating process. They argue that sales contracts are more primitive than cash receipts and the modeling of the sales process should precede the modeling of the earnings and cash flows processes. The fundamental analysis of security valuation that uses signals related to inventory, gross margin, receivables, SG&A, order backlog, and labor efficiency that are all benchmarked against sales performance (Lev and Thiagarajan 1993) is also consistent with this sales-centric view of evaluating economic performance.

---

[5] As discussed in Kothari (2001), one motivation for studying the properties of earnings components is to assess their incremental information beyond earnings in their association with stock prices. Earnings forecast models based on earnings components are often implicitly assumed in such research contexts, even though some studies may not explicitly examine these models of predictive ability (e.g., Lipe 1986).

Cost Variability and Cost Stickiness (CVCS) Model

We consider accounting earnings (E_t) in period t, measured as sales revenue (S_t) net of costs (C_t):[6]

$$E_t = S_t - C_t \tag{1}$$

Management accounting textbooks describe two basic types of cost behavior patterns: variable costs and fixed costs (Garrison and Noreen 2002). Variable costs are defined as those that change in proportion to changes in sales volume, whereas fixed costs are characterized as those that remain unchanged in a relevant range. From an economic perspective, variable costs represent the flexible productive resources that managers can easily adjust in producing goods and services, whereas fixed costs represent the committed resources invested to provide long-term productive capacity and thus are not expected to change with short-term production volume. Such an economic interpretation of different earnings components serves as a conceptual basis for modeling the earnings-generating process conditional on the relationship between sales and costs. Using this cost model with sales revenue as the cost driver, we express total costs in period t as:

$$C_t = vS_t + F_t \tag{2}$$

where v represents the constant unit variable cost.

It follows from Equations (1) and (2) that:

$$E_t = (1 - v)S_t - F_t \tag{3}$$

We assume that S_t follows a first order autoregressive process and allow the AR(1) coefficient to depend on whether sales decrease in period $t-1$:[7]

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \varepsilon_S \tag{4}$$

where parameter α_1 represents the persistence of sales.

Anderson et al. (2003) document that costs are sticky in that they decrease less with a sales decrease than they increase with a sales increase. Such sticky cost behavior arises if trimming off excess resources when demand declines is relatively more costly than scaling up resources to accommodate increased demand. Under such conditions, managers' rational

---

[6] We use S_t to denote sales in dollars instead of volume, and we interpret v as the fraction of sales representing variable costs. Our interpretation is consistent with our empirical forecast model that uses data on sales revenue. Since volume data is unavailable, our study is limited to use of revenue data, which confounds price, volume, and mix changes.

[7] We are grateful to the referees for suggesting alternative specifications for the sales process that allow the AR(1) coefficient to depend on whether sales decrease in period $t-1$ as in the following two equations:

$$S_t = \alpha_0 + \alpha_1 S_{t-1} + \alpha_2 S_{t-1} D_{t-1} + \varepsilon_S$$

and $\Delta S_t = \alpha_0 + \alpha_1 \Delta S_{t-1} + \alpha_2 \Delta S_{t-1} D_{t-1} + \varepsilon_S$

where $D_{t-1} = 1$ if sales decline in period $t-1$, and 0 otherwise. The parameter α_2 captures the possibly asymmetric behavior in sales changes. To address the concern that such asymmetric sales behavior confounds our results, we estimated models based on the above two alternative specifications using a pooled sample in 1992–2002 and also on a yearly basis. Our results are robust to these alternative specifications, indicating that the possibly asymmetric sales changes are not a confounding factor in the forecast performance of our CVCS model.

resource adjustment decision entails less reduction of resources when sales decrease than addition of resources when sales increase (Anderson et al. 2003). Unlike the *static* traditional cost model that relates costs only to the contemporaneous level of sales volume independent of costs and volume in the prior period, cost stickiness exemplifies *dynamic* cost behavior that depends on costs and sales in the prior period and the direction of change in sales from the prior period.

We model fixed costs as a first-order autoregressive process specified as:

$$F_t = \beta_0 + \beta_1 F_{t-1} + \beta_2 v(S_{t-1} - S_t)D_t + \varepsilon_F \tag{5}$$

where $D_t = 1$ if sales decline in period t, and 0 otherwise. The term $\beta_2 v(S_{t-1} - S_t)D_t$ in Equation (5) represents the additional amount of costs incurred due to costs being sticky when sales decrease in period t. That is, a portion β_2 of the variable costs in the amount of $v(S_{t-1} - S_t)D_t$ is no longer driven down by a decrease in sales and as a result resembles a "fixed" cost. Costs that may exhibit such "sticky" behavior include skilled labor payroll costs, advertising and sales promotion costs, and branch operating costs.[8] We recognize cost stickiness in the context of our variable and fixed cost model by allowing the level of fixed costs, F_t, to increase when sales decline.

Substituting S_t and F_t in Equation (3) with Equations (4) and (5), we have:

$$\begin{aligned}
E_t &= (1 - v)S_t - F_t \\
&= (1 - v)(\alpha_0 + \alpha_1 S_{t-1} + \varepsilon_S) - (\beta_0 + \beta_1 F_{t-1} + v\beta_2(S_{t-1} - S_t)D_t + \varepsilon_F) \\
&= (1 - v)(\alpha_0 + \alpha_1 S_{t-1} - \varepsilon_S) - \beta_0 - \beta_1((1 - v)S_{t-1} - E_{t-1}) \\
&\quad - v\beta_2(-\alpha_0 + (1 - \alpha_1)S_{t-1} - \varepsilon_S)D_t - \varepsilon_F \\
&= (1 - v)\alpha_0 - \beta_0 + \beta_1 E_{t-1} + (1 - v)(\alpha_1 - \beta_1)S_{t-1} - v\beta_2(1 - \alpha_1)S_{t-1}D_t \\
&\quad + v\beta_2 D_t \varepsilon_S - \varepsilon_F \\
&= \gamma_0 + \gamma_1 D_t + \gamma_2 E_{t-1} + \gamma_3 S_{t-1} + \gamma_4 S_{t-1}D_t + \eta
\end{aligned} \tag{6}$$

where:

$\gamma_0 = (1 - v)\alpha_0 - \beta_0$;
$\gamma_1 = v\alpha_0\beta_2$;
$\gamma_2 = \beta_1$;
$\gamma_3 = (1 - v)(\alpha_1 - \beta_1)$;
$\gamma_4 = v(\alpha_1 - 1)\beta_2$; and
$\eta = v\beta_2 D_t \varepsilon_S - \varepsilon_F$.

---

[8] Resources that are not mechanically linked to overall activity levels are likely to exhibit sticky behavior. For example, the payroll costs of contract employees could be reasonably viewed as variable with total revenue. However, in a period of sales downturn perceived to be temporary, managers may choose not to reduce the workforce proportionately in order to avoid having to incur additional costs of recruiting and retaining them when sales levels pick up in a subsequent period. Payroll costs for such employees retained in excess of the levels required for the reduced sales activity resemble "fixed costs" more than "variable costs" as traditionally defined.

Unlike E_{t-1}, S_{t-1}, and D_{t-1}, which are all observed at the beginning of period t, the sales decrease dummy variable D_t needs to be estimated before earnings E_t can be forecasted using model (6).[9] Letting $\hat{D}_t$ denote the estimated value of D_t based on information available at the beginning of period t, $\hat{E}_t$, the forecasted earnings for period t, is given by:

$$\hat{E}_t = \gamma_0 + \gamma_1\hat{D}_t + \gamma_2E_{t-1} + \gamma_3S_{t-1} + \gamma_4S_{t-1}\hat{D}_t \qquad (7)$$

The forecast model is expressed in terms of observed earnings and sales from prior years. Fixed and variable costs need not to be observed or estimated separately.

III. SAMPLE DATA AND MODEL ESTIMATION

Data

Our sample period spans 1988 through 2002.[10] All four time-series forecast models we consider require information on earnings or components of earnings based on line items from the income statement and the cash flows statement. These annual financial statement data are obtained from the Compustat database. We define earnings as income before extraordinary items (#18) and compute return on owners' equity (*ROE*) by dividing earnings by beginning book value of equity (#60). We follow Fairfield et al. (1996) to measure the operating income (*OPINC*) and non-operating income (*NOPTAX*) components of earnings. Specifically, *OPINC* is measured as operating income after depreciation (#178) net of interest expense (#15), special items (#17), and minority interest (#49). *NOPTAX* consists of non-operating income (#61) net of income taxes (#16). Consistent with the accounting accruals literature (Collins and Hribar 2000), we measure the cash flows component (*CFO*) of earnings based on SFAS No. 95 data reported in the cash flows statement (#308 − #124) and the accruals component (*ACCRUALS*) as the difference between earnings and cash flows (#18 − #308 + #124).[11] Our CVCS model requires, in addition to lagged ROE, information about sales revenue (SALES), which we measure as net sales revenue (#12). *OPINC*, *NOPTAX*, *CFO*, *ACCRUALS*, and *SALES* are all scaled by the beginning-of-year owners' equity. We compute analysts' consensus forecasts of one-year-ahead ROE comparable to forecasts generated from the time-series model as the mean analysts' forecasts of earnings per share (EPS) reported on the I/B/E/S database nine months before fiscal year-end, divided by beginning-of-year book value of equity per share.[12]

To mitigate the potential adverse effect of data errors and outliers on the estimation of the time-series models and on the evaluation of forecast accuracy of these models and analysts' forecasts, we restrict our sample to firm-year observations with (1) positive values for owners' equity; (2) absolute values of ROE and lagged ROE less than 1; and (3) absolute values of net profit margin (i.e., the ratio of earnings to sales revenue) less than 1.[13] Since we estimate the time-series forecast models using industry-level (based on four-digit SIC code) random coefficient regressions (discussed in detail in the next subsection), we delete firms from each forecast year and the associated estimation period (i.e., the four years immediately prior to the forecast year) that are in industries with less than 20 firm-year observations available in the estimation period. In order to calculate the forecast errors in ROE from all four time-series models and compare their forecast accuracy performance simultaneously, we require all sample observations to have no missing values for ROE in each forecast year and for the lagged variables: *ROE*, *OPINC*, *NOPTAX*, *CFO*, *ACCRUALS*, and *SALES*.

---

[9] The actual value of the sales decrease dummy variable in year t is not available when constructing earnings forecast for year t so we cannot use it in any earnings forecast model.

[10] We begin our study period from 1988 because SFAS No. 95 data required for computing cash flows from continuing operations (CFO) were not available prior to 1988. The data on CFO is used to decompose earnings into cash flows and accruals components in the CASHFLOW model (Collins and Hribar 2000).

[11] We require the earnings data reported in the statement of cash flows (#123) to be the same as that reported in the income statement (#18) to ensure comparability of cash flows/accruals decomposition of earnings with the operating income/non-operating income decomposition.

[12] This particular measure of analysts' forecasts of one-year-ahead ROE is comparable to forecasts generated from the time-series models using lagged annual data in the sense that the timing of the consensus forecasts roughly coincides with the availability of previous fiscal year's financial statement information. Our method of deriving analysts' forecasts of one-year-ahead ROE from their EPS forecasts follows that of Frankel and Lee (1998).

These sample selection procedures result in a full sample of 39,367 firm-year observations from 8,771 firms over the 11-year forecast period 1992–2002. To evaluate forecast accuracy of the four time-series models relative to that of analysts, we obtain a subsample of 15,500 firm-year observations from 4,334 firms with available analysts' *ROE* forecasts. Descriptive statistics on *ROE* and components of *ROE* for the full sample are reported in Table 1. The median (mean) *ROE* is 0.092 (0.052) and the median (mean) *SALES* is 2.025 (3.021). The median (mean) *OPINC* is 0.109 (0.079), the median (mean) *NOPTAX* is −0.019 (−0.027), the median (mean) *CFO* is 0.175 (0.174) and the median (mean) *ACCRUALS* is −0.106 (−0.122).

TABLE 1
Descriptive Statistics on Return on Equity (ROE) and Components of ROE
Sample Period: 1992–2002 (n = 39,367)

Variable	Median	Mean	Std. Dev.	Q1	Q3
ROE	0.092	0.052	0.250	−0.022	0.173
SALES	2.025	3.021	5.977	1.067	3.677
OPINC	0.109	0.079	0.325	−0.045	0.239
NOPTAX	−0.019	−0.027	0.152	−0.077	0.017
CFO	0.175	0.174	0.517	0.031	0.318
ACCRUALS	−0.106	−0.122	0.505	−0.237	0.000

ROE = return on equity, measured as net income before extraordinary items (Compustat #18) divided by beginning-of-year book value of owners' equity (#60);
SALES = net sales revenue (#12) divided by beginning-of-year book value of owners' equity;
OPINC = operating income (#178 − #15 + #17 − #49) divided by beginning-of-year book value of owners' equity;
NOPTAX = non-operating income (#61) net of income taxes (#16) divided by beginning-of-year book value of owners' equity;
CFO = cash flows from continuing operations (#308 − #124) divided by beginning-of-year book value of owners' equity; and
ACCRUALS = total accruals (#18 − #308 + #124) divided by beginning-of-year book value of owners' equity.

---

[13] Our sample screening procedure consists of two steps. The first step screens each *estimation sample* based on all information available up to the end of estimation period (i.e., year $t-4$ to year $t-1$ in our case). This step (estimation and forecasting) is implementable *ex ante* since no information beyond year $t-1$ is used to cull the sample. The second step screens the forecast sample by deleting extreme values of ROE in the forecast year (i.e., year t) to evaluate the predictive performance of competing forecast models. This evaluation is *ex post* since it requires *actual values* of ROE in year t. This step helps alleviate concerns about outliers possibly inflating the average absolute forecast errors and thus obscuring the differential predictive ability of competing forecast models. Our results about the superior predictive ability of our CVCS model hold even when we do not employ any screening of the forecast sample.

Forecast Model Estimation

We benchmark the performance of our CVCS model in forecasting *ROE* against a simple model based only on lagged *ROE* data (the ROE model), the OPINC model based on lagged operating and non-operating income disaggregation, and the CASHFLOW model based on lagged cash flows and accruals components. Specifically, the four time-series ROE forecast models we consider are:

ROE model: $$ROE_t = \gamma_{a0} + \gamma_{a1}ROE_{t-1} + \varepsilon_t \qquad (8)$$

OPINC model: $$ROE_t = \gamma_{b0} + \gamma_{b1}OPINC_{t-1} + \gamma_{b2}NOPTAX_{t-1} + \varepsilon_t \qquad (9)$$

CASHFLOW model: $$ROE_t = \gamma_{c0} + \gamma_{c1}CFO_{t-1} + \gamma_{c2}ACCRUALS_{t-1} + \varepsilon_t \qquad (10)$$

CVCS model: $$\begin{aligned} ROE_t = {} & \gamma_{d0} + \gamma_{d1}DECRDUM_t + \gamma_{d2}ROE_{t-1} \\ & + \gamma_d 3SALES_{t-1} + \gamma_{d4}SALES_{t-1}DECRDUM_t \\ & + \varepsilon_t \end{aligned} \qquad (11)$$

where $DECRDUM_t$ is a sales decrease indicator variable that takes on the value 1 when the unscaled sales revenue in period t is less than in period $t-1$, and 0 otherwise.

We employ industry-level (based on four-digit SIC code) random coefficient regressions (RCR) to estimate our CVCS model.[14] We allow for stochastic heterogeneity in model parameters for firms in each industry. The coefficients for the firms are random variates drawn from a separate distribution (mean and variance) for each industry. To illustrate, the set of coefficients in Equation (11) is assumed to have the following underlying stochastic structure:

$$\gamma_{di} = \gamma_{di}^{IND} + \delta_{di}^{FIRM},\ i = 0, 1, \ldots, 4$$

where γ_{di}^{IND} is the mean (constant) for a given industry in which a particular firm belongs and $\delta_{di}^{FIRM} \sim N(0, \sigma_{IND}^2)$ is the firm's deviation from the industry mean. Each δ_{di}^{FIRM} follows a normal distribution with mean 0 and industry-specific variance σ_{IND}^2. Since we use a pooled sample, there remains residual unexplained variation ε_t in ROE_t for each firm in any period t. Similar stochastic structure also applies to coefficients in the other three times-series models (i.e., Equations (8)–(10)), allowing us to estimate those coefficients using RCR. We do this to ensure that the four alternative models are evaluated on a comparable basis.

Following Fairfield et al. (1996), we estimate the parameters of all four forecast models for each year on a rolling basis using all available observations from the preceding four-year period.[15] For example, we use data from 1998–2001 to estimate the forecast model parameters applicable to constructing the ROE forecast for 2002; we obtain a new set of parameter estimates using data from 1997–2000 in order to derive the ROE forecast for 2001, and so on. Since we consider an 11-year forecast period from 1992–2002, we have 11 corresponding overlapping four-year estimation periods from 1988–2001. This industry-level RCR approach produces, for each model, coefficient estimates of the industry mean and deviation of each individual firm from its industry mean. The estimate of a firm-specific random-effect coefficient is given by the sum of the industry mean and firm-specific deviation.[16] We use this firm-specific estimate to compute the *ROE* forecast during the forecast period. To illustrate in algebraic terms, the set of RCR coefficient estimates in Equation (11) is expressed as:

---

[14] Firm-specific estimation of our CVCS model is not feasible because the model has seven parameters to be estimated and this requires a long time-series of data if firm-level (rather than pooled) estimation is conducted. However, the assumption that model parameters (e.g., cost structure) are constant over time is also difficult to sustain when a long time-series of data is employed to ensure a reasonable number of degrees of freedom for the estimation. Therefore, the compromise we adopt is industry-level estimation with random coefficients for firms in each industry.

[15] This rolling estimation approach is also used in Fairfield et al. (1996). Our method differs from theirs in that we use industry-level random coefficient regressions based on four-year estimation period, whereas they use pooled regressions based on seven-year estimation period.

$$\hat{\gamma}_{di} = \hat{\gamma}_{di}^{IND} + \hat{\delta}_{di}^{FIRM}, \; i = 0, 1, \ldots, 4$$

where $\hat{\gamma}_{di}^{IND}$ is the estimated industry mean and $\hat{\delta}_{di}^{FIRM}$ is an individual firm's estimated deviation from the industry mean.

Table 2 reports the means and standard deviations of estimated coefficients of the time-series forecast models calculated over the 11 four-year rolling industry-level (based on four-digit SIC code) random coefficient regressions.[17] For the ROE model, the average slope coefficient on ROE_{t-1} is 0.513. The average slope coefficient estimates on $OPINC_{t-1}$ and $NOPTAX_{t-1}$ in the OPINC model are 0.485 and 0.352, respectively. The different slope coefficient estimates on $OPINC_{t-1}$ and $NOPTAX_{t-1}$ suggest that decomposing earnings into operating and non-operating income may provide incremental predictive power for forecasting one-year-ahead ROE. Similarly, the slope coefficients on the cash flows (CFO_{t-1}) and accruals ($ACCRUALS_{t-1}$) components in the CASHFLOW model are different (0.541 and 0.473, respectively).[18]

IV. FORECAST ACCURACY COMPARISONS

Main Results

To forecast *ROE* with our CVCS model, we require in addition to the parameter estimates an indicator of whether sales increase or decrease in the forecast year. This estimator must be based on information available prior to the forecast period. We estimate the sales decrease indicator $DECRDUM_t$ using a logit regression of the following form:

$$\text{Prob}(DECRDUM_t = 1) = \frac{e^y}{1 + e^y} \tag{12}$$

---

[16] In addition to being firm-specific, the gammas in Equations (8)–(11) and (14)–(17) are indexed by estimation periods because of our rolling estimation method. We omit both firm and time indices in these equations to avoid cumbersome notation.

[17] The summary statistics reported in Table 2 are based on estimated coefficients for individual firms retrieved from the estimation of the random coefficient regression (RCR) model averaged over 8,771 firms in all 432 industries and over the 11 estimation periods. The mean and standard deviation reported in Table 2 are the mean and standard deviation of those coefficient estimates.

[18] Sloan (1996) estimated a similar regression model of earnings on lagged cash flows and accruals. Using pooled estimation, he reported 0.855 and 0.765 as the coefficients on the lagged cash flows and accruals, respectively, and using industry-level estimation, he reported 0.781 and 0.721 as the means of these coefficients. Our results are consistent with Sloan's to the extent that the coefficient estimate on cash flows is greater than that on accruals. Sloan's estimated magnitude of these coefficients may differ from ours because he (1) used a different sample period (from 1962 to 1991); (2) scaled earnings, cash flows, and accruals by average total assets; and (3) used pooled (or industry-level pooled) regressions.

TABLE 2
Coefficient Estimates[a] of Time-Series ROE Forecast Models from 11 Four-Year Rolling Industry-Level (four-digit SIC) Random Coefficient Regressions
Sample Estimation Period: 1988–2001

	ROE Model	OPINC Model	CASHFLOW Model	Asy. ROE Model	CVCS Model
Independent Variable	**Mean Coeff. Est. (Std. Dev.)**	**Mean Coeff. Est. (Std. Dev.)**	**Mean Coeff. Est. (Std. Dev.)**	**Mean Coeff. Est. (Std. Dev.)**	**Mean Coeff. Est. (Std. Dev.)**
Intercept	0.015 (0.028)	0.013 (0.028)	0.007 (0.025)	0.030 (0.021)	0.042 (0.020)
$DECRDUM_t$					−0.088 (0.043)
$LOSS$				−0.054 (0.029)	
ROE_{t-1}	0.513 (0.094)			0.464 (0.064)	0.475 (0.095)
$LOSS * ROE_{t-1}$				−0.070 (0.106)	
$OPINC_{t-1}$		0.485 (0.065)			
$NOPTAX_{t-1}$		0.352 (0.104)			
CFO_{t-1}			0.541 (0.058)		
$ACCRUALS_{t-1}$			0.473 (0.071)		
$SALES_{t-1}$					0.002 (0.004)
$SALES_{t-1} * DECRDUM_t$					−0.008 (0.008)
Average Adj. R^{2b}	0.2649	0.2911	0.3030	0.2985	0.3676

---

[a] The numbers reported in each cell are the mean and standard deviation (in parentheses) of firm- and period-specific coefficients estimated using the random coefficient regression (RCR) over 8,771 firms in 432 industries and over the 11-year estimation period.

[b] Since a goodness-of-fit measure like R^2 is not well-defined in the context of RCR estimation, the adjusted R^2s are obtained from estimating the forecasting models using OLS for each four-digit-SIC industry and each estimation period. The average is calculated across all industries and estimation periods.

ROE model: $ROE_t = \gamma_{a0} + \gamma_{a1} ROE_{t-1} + \varepsilon_{at}$

OPINC model: $ROE_t = \gamma_{b0} + \gamma_{b1} OPINC_{t-1} + \gamma_{b2} NOPTAX_{t-1} + \varepsilon_{bt}$

CASHFLOW model: $ROE_t = \gamma_{c0} + \gamma_{c1} CFO_{t-1} + \gamma_{c2} ACCRUALS_{t-1} + \varepsilon_{ct}$

Asy. ROE model: $ROE_t = \gamma_{e0} + \gamma_{e1} LOSS + \gamma_{e2} ROE_{t-1} + \gamma_{e3} LOSS{*}ROE_{t-1} + \varepsilon_t$

CVCS model: $ROE_t = \gamma_{d0} + \gamma_{d1} DECRDUM_t + \gamma_{d2} ROE_{t-1} + \gamma_{d3} SALES_{t-1} + \gamma_{d4} SALES_{t-1} DECRDUM_t + \varepsilon_{dt}$

(continued on next page)

TABLE 2 (Continued)

ROE = return on equity, measured as net income before extraordinary items divided by beginning-of-year book value of owners' equity;
OPINC = operating income divided by the beginning-of-year owners' equity;
NOPTAX = non-operating income net of income taxes divided by beginning-of-year book value of owners' equity;
CFO = cash flows from continuing operations divided by beginning-of-year book value of owners' equity;
ACCRUALS = total accruals divided by beginning-of-year book value of owners' equity;
DECRDUM = a dummy variable that takes the value 1 if sales revenue decreases from prior year, and 0 otherwise;
LOSS = a dummy variable that takes the value 1 if earnings in the prior year are negative, and 0 otherwise; and
SALES = net sales revenue divided by beginning-of-year book value of owners' equity.

where $y = \beta_0 + \beta_1 \Delta REV_{t-1} + \varepsilon$, ΔREV_{t-1} is the percentage change in sales revenue from year $t-2$ to year $t-1$. We estimate the parameters β_0 and β_1 of this logit model on a rolling basis for each forecast year using pooled data from the preceding four-year period.[19] Table 3 shows the estimation results of logit regressions. The estimated mean coefficient on ΔREV_{t-1} across the 11 estimation periods is $\hat{\beta}_1 = -1.865$, and the corresponding mean odds ratio estimate is 0.157. This result indicates that a one percentage point decrease in sales in year $t-1$ is associated with a six-fold (i.e., $1/0.157 \approx 6$) increase in the odds ratio

TABLE 3
Estimation Results of Logit Regression of Sales Decrease Dummy on the Lagged Percentage Change in Sales Revenue From 11 Four-Year Rolling Regressions
Sample Estimation Period: 1988–2001

$$\text{Logit Model: } DECRDUM_t = \frac{\exp(\beta_0 + \beta_1 \Delta REV_{t-1} + \varepsilon_t)}{1 + \exp(\beta_0 + \beta_1 \Delta REV_{t-1} + \varepsilon_t)}$$

Panel A: Coefficient Estimates and Odds Ratio Estimates

Independent Variable	Mean Coeff. Est. (Std. Dev.)	Mean Odds Ratio (Std. Dev.)	Mean Lower Confid. Limit (Std. Dev.)	Mean Upper Confid. Limit (Std. Dev.)
Intercept	−0.743 (0.144)			
ΔREV_{t-1}	−1.865 (0.155)	0.157 (0.024)	0.133 (0.020)	0.185 (0.029)

Panel B: Measures of Association

Mean Percent Concordant (Std. Dev.)	Mean Percent Discordant (Std. Dev.)	Mean Percent Tied (Std. Dev.)
64.190 (0.717)	35.080 (0.704)	0.730 (0.053)

$DECRDUM_t$ = a dummy variable that takes the value 1 if sales revenue decreases in year t, and 0 otherwise; and
ΔRET_{t-1} = percentage change in sales revenue from year $t-2$ to year $t-1$.

---

[19] To mitigate concerns for undue influence of potential outliers, we deleted observations with absolute value of ΔREV_{t-1} greater than 1 in each round of estimation.

of sales decline in year *t*. The high concordant measure of 64 percent indicates that this simple and parsimonious logit model (12) performs reasonably well in predicting the direction of change in sales revenue.

We determine the estimated value of the sales decrease dummy variable ($\overline{DECRDUM_t} \equiv \hat{D}_t$) for the forecast period as:

$$\overline{DECRDUM_t} = \begin{cases} 1, & \text{if } \hat{\beta}_0 + \hat{\beta}_1 \Delta REV_{t-1} > 0 \\ 0, & \text{if } \hat{\beta}_0 + \hat{\beta}_1 \Delta REV_{t-1} \leq 0 \end{cases} \tag{13}$$

Our classification scheme based on the zero cutoff in Equation (13) is interpreted as a simple heuristic that predicts a sales decline ($DECRDUM = 1$) if the estimated probability is greater than 50 percent.

The forecast of *ROE* from our CVCS model is then given by:

$$\begin{aligned} \overline{ROE_t} = {} & \hat{\gamma}_{d0} + \hat{\gamma}_{d1}\overline{DECRDUM_t} + \hat{\gamma}_{d2}ROE_{t-1} + \hat{\gamma}_{d3}SALES_{t-1} \\ & + \hat{\gamma}_{d4}SALES_{t-1}\overline{DECRDUM_t} \end{aligned} \tag{14}$$

where, $\hat{\gamma}_{d0}, \hat{\gamma}_{d1},\ldots,\hat{\gamma}_{d4}$ are the five estimated coefficients of the CVCS model in Equation (11).

The *ROE* forecasts from the ROE model, the OPINC model, and the CASHFLOW model, respectively, are as follows:

$$\overline{ROE_t} = \hat{\gamma}_{a0} + \hat{\gamma}_{a1}ROE_{t-1} \tag{15}$$

$$\overline{ROE_t} = \hat{\gamma}_{b0} + \hat{\gamma}_{b1}OPINC_{t-1} + \hat{\gamma}_{b2}NOPTAX_{t-1} \tag{16}$$

$$\overline{ROE_t} = \hat{\gamma}_{c0} + \hat{\gamma}_{c1}CFO_{t-1} + \hat{\gamma}_{c2}ACCRUALS_{t-1} \tag{17}$$

where $\hat{\gamma}$s in the above expressions are the estimated coefficients of their corresponding models in Equations (8), (9), and (10).

Based on the estimated parameters of the four forecast models and the estimated direction of change in sales revenue, we obtain four forecasts of *ROE* in forecast period *t* ($\overline{ROE_t}$) as given previously in Equations (14)–(17). The forecast error (*FE*) is defined as the difference between the actual realization of *ROE* in year *t* and the forecasted $\overline{ROE_t}$:

$$FE_t = ROE_t - \overline{ROE_t} \tag{18}$$

The absolute forecast error (*AFE*) is the absolute value of the forecast error *FE*:

$$AFE_t = |FE_t| \tag{19}$$

The statistics on the distribution of *AFE* for the time-series models are shown in Panel A of Table 4. The CVCS model generates *AFE* with the smallest median (0.0734), mean (0.1412), first quartile (0.0291), and third quartile (0.1779) of the four time-series models considered. In contrast, the basic ROE model produces *AFE* with the largest median (0.0805), mean (0.1427), first quartile (0.0340), and third quartile (0.1790). The OPINC and CASHFLOW models represent intermediate cases in terms of their forecast accuracy

TABLE 4
ROE Forecast Accuracy Comparisons among Time-Series Forecast Models Estimated From 11 Four-Year Rolling Industry-Level (four-digit SIC) Random Coefficient Regressions
Sample Forecast Period: 1992–2002 (n = 39,367)

Panel A: Distributions of Absolute Forecast Errors (AFE)

Forecast Model	Median	Mean	Std. Dev.	Q1	Q3
ROE	0.0805	0.1427	0.1735	0.0340	0.1790
OPINC	0.0797	0.1422	0.1736	0.0334	0.1784
CASHFLOW	0.0797	0.1420	0.1740	0.0337	0.1780
Asy. ROE	0.0791	0.1423	0.1738	0.0328	0.1784
CVCS	0.0734	0.1412	0.1788	0.0291	0.1779

Panel B: Distributions of Pair-Wise Differences in Absolute Forecast Errors (AFE)[a]

Base Model	Comparison Model	Median Difference	Mean Difference	Std. Dev. of Difference
ROE	OPINC	0.0002***	0.0005^{+++}	0.0215
ROE	CASHFLOW	0.0005***	0.0006^{+++}	0.0370
ROE	CVCS	0.0052***	0.0010^{+++}	0.0446
OPINC	CASHFLOW	0.0001	0.0001	0.0400
OPINC	CVCS	0.0035***	0.0009^{+++}	0.0431
CASHFLOW	CVCS	0.0031***	0.0008^{+++}	0.0493

***, **, * indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, using Wilcoxon signed rank test.

$^{+++}$, $^{++}$, $^{+}$ indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, using the t-test.

[a] Pair-wise difference in AFE = AFE from the base model − AFE from the comparison model.

measured using distributions of *AFE*. The forecast accuracy improvement of the CVCS model relative to the benchmark ROE model in terms of reduction in median *AFE* is 0.0071 (=0.0805 − 0.0734), which is about nine times as large as the improvement of the OPINC or CASHFLOW models relative to the ROE model (reduction in median *AFE* = 0.0008).

To provide a statistical test of the forecast improvement, following Fairfield et al. (1996), we employ the nonparametric Wilcoxon signed rank test to evaluate the paired differences in *AFE* between two forecast models without imposing restrictive assumptions about the exact form of the underlying distribution (Lehmann 1975; Sheskin 1997).[20] Panel B of Table 4 presents pair-wise comparisons of *AFE* among the four models based on the distribution of the difference in *AFE* between any two competing models. Included in each row are the median, mean, and standard deviation of the difference in *AFE* between a comparison model and a base model. Positive (negative) values of the difference in *AFE* indicate that the comparison model generates a lower (higher) level of *AFE* and thus is more (less) accurate in forecasting one-year-ahead *ROE*, as compared to the base model. All the results based on pair-wise comparisons are consistent with those based on comparisons at the median or mean *AFE* in Panel A of Table 4. The median reduction in *AFE* of the OPINC model relative to the ROE model is 0.0002 ($p < 0.001$).[21] The CASHFLOW model also provides more accurate forecasts than the ROE model, with a median improvement of 0.0005 ($p < 0.001$). In contrast to the OPINC and CASHFLOW models that exhibit modest forecast accuracy improvement over the ROE model, the CVCS model improves forecast accuracy substantially. The median reductions in the *AFE* of CVCS relative to the ROE, OPINC, and CASHFLOW models are 0.0052 ($p < 0.001$), 0.0035 ($p < 0.001$), and 0.0031 ($p < 0.001$), respectively. Examining the mean differences in *AFE* produces qualitatively similar results.[22]

---

[20] Since the hold-out samples used in comparing any two competing forecast models consist of the same set of firm-year observations, the appropriate statistical test of relative forecast accuracy is based on a matched sample comparison of AFEs.

The Table 4 evidence confirms the previous result that disaggregating earnings into its operating and non-operating components leads to more accurate forecasting of one-year-ahead ROE. It also demonstrates that decomposing earnings into cash flows and accruals has a similar impact on ROE forecast accuracy. More importantly, decomposing earnings based on cost variability and cost stickiness provides substantial incremental predictive ability in forecasting future ROE, over and above the forecast improvement of earnings decomposition based on financial statement line items that do not recognize the relation between sales and costs.

The Table 5 results are based on a subsample of firms containing analysts' consensus forecasts. Using the Frankel and Lee (1998) procedure to convert earnings per share (EPS) forecasts to ROE forecasts, we calculate the analysts' *AFE* in ROE based on their EPS forecasts reported on the I/B/E/S database. The statistics on the *AFE* distributions of analysts' forecasts and time-series model forecasts are presented in Panel A of Table 5. Consistent with evidence of analyst earnings forecast superiority *vis-à-vis* time-series models (Brown 1993), the analysts' consensus forecasts are the most accurate, with median (mean) *AFE* of 0.0457 (0.0976), which is 59 percent (80 percent) of the median (mean) *AFE* from the benchmark ROE model (median $AFE = 0.0775$, mean $AFE = 0.1215$). Consistent with the results in Table 4, the CVCS model is the most accurate of the three time-series models based on lagged earnings components, with median (mean) *AFE* of 0.0675 (0.1168). The OPINC and the CASHFLOW models have very similar forecasting performance, with median (mean) *AFE* of 0.759 (0.1205) and 0.760 (0.1206), respectively.

Panel B of Table 5 reports the results of comparing forecast accuracy based on pair-wise differences in *AFE*. The median reductions in matched *AFE* from the OPINC and the CASHFLOW models relative to the ROE model are 0.0005 and 0.0006, respectively, which is less than 3 percent of the median reduction in matched AFE from I/B/E/S consensus forecasts (0.0206). In contrast, the median reduction in matched *AFE* from our CVCS model is 0.0113, over half (55 percent) of that attributable to analysts' consensus forecasts. Similar to the Table 4 findings, these results indicate that the OPINC and CASHFLOW models are almost indistinguishable, with median difference of −0.0002 and mean difference of 0.0001. Both differences are statistically significant at 1 percent. In addition, I/B/E/S forecasts beat both the OPINC and CASHFLOW models and, to a lesser extent, the CVCS model, in pair-wise comparisons of forecast accuracy.

To assess the significance of forecast improvement of the three models of earnings decomposition, we calculate the reduction in median and mean *AFE* from the OPINC,

---

[21] Fairfield et al. (1996) reported a median reduction of 0.0008 in absolute forecast error using the OPINC model relative to the ROE model.

[22] We conducted a *t*-test for the equality of the means of two related samples. The t-statistics are all significant at the 0.001 level, indicating that both the OPINC and the CASHFLOW models perform better than the ROE model but worse than the CVCS model. The two-sample *t*-test requires the assumption of normal distribution in both samples. Since the normality assumption is found not to be descriptive of the distributions of AFE, we view the *t*-test results as suggestive, not conclusive.

TABLE 5
ROE Forecast Accuracy Comparison between Time-Series Forecast Models and I/B/E/S Consensus Forecasts[a]
Sample Forecast Period: 1992–2002 (n = 15,500)

Panel A: Distributions of Absolute Forecast Errors (AFE)

Forecast Model	Median	Mean	Std. Dev.	Q1	Q3
ROE	0.0775	0.1215	0.1364	0.0364	0.1511
OPINC	0.0759	0.1205	0.1370	0.0359	0.1490
CASHFLOW	0.0760	0.1206	0.1374	0.0354	0.1499
Asy. ROE	0.0747	0.1202	0.1369	0.0345	0.1511
CVCS	0.0675	0.1168	0.1412	0.0289	0.1454
I/B/E/S	0.0457	0.0976	0.1388	0.0169	0.1156

Panel B: Distributions of Pair-Wise Differences in Absolute Forecast Errors (AFE)[c]

Base Model	Comparison Model	Median Difference	Mean Difference	Std. Dev. of Difference
ROE	OPINC	0.0005***	0.0008^{+++}	0.0183
ROE	CASHFLOW	0.0006***	0.0009^{+++}	0.0309
ROE	CVCS	0.0113***	0.0047^{+++}	0.0354
ROE	I/B/E/S	0.0206***	0.0239^{+++}	0.1148
OPINC	CASHFLOW	−0.0002	0.0001	0.0335
OPINC	CVCS	0.0075***	0.0037^{+++}	0.0387
OPINC	I/B/E/S	0.0190***	0.0229^{+++}	0.1153
CASHFLOW	CVCS	0.0075***	0.0038^{+++}	0.0437
CASHFLOW	I/B/E/S	0.0230***	0.0230^{+++}	0.1158
CVCS	I/B/E/S	0.0105***	0.0192^{+++}	0.1049

Panel C: Forecast Improvement over the ROE Model[b]

	Accuracy Improvement over the ROE Model		% of Accuracy Improvement of I/B/E/S over the ROE Model	
Forecast Model	Difference in Median AFE	Difference in Mean AFE	Difference in Median AFE	Difference in Mean AFE
OPINC	0.0016	0.0010	5.0	4.2
CASHFLOW	0.0015	0.0009	4.7	3.8
CVCS	0.0100	0.0047	31.4	19.7
I/B/E/S	0.0318	0.0239	100.0	100.0

***, **, * indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, using Wilcoxon signed rank test.

$^{+++}$, $^{++}$, $^{+}$ indicate significance at the 1 percent, 5 percent, and 10 percent levels, respectively, using the t-test.

[a] ROE forecasts calculated based on analysts' consensus forecasts of earnings per share reported on I/B/E/S database nine months before the fiscal year end.

[b] Forecast improvement is measured as the difference in median or mean absolute forecast errors between two forecast methods.

[c] Pair-wise difference in AFE = AFE from the base model − AFE from the comparison model.

CASHFLOW, and CVCS models, respectively, as a percentage of the reduction in *AFE* of analysts' forecasts, all benchmarked against the ROE model. The differential forecast accuracy between analysts' forecasts and the ROE model reflects the total "information gap" existing between the basic time-series model that uses only aggregate historical earnings information and financial analysts who utilize a significantly richer set of information, possibly including information contained in earnings components, in making earnings forecasts. The results are shown in Panel C of Table 5. The CVCS model improves forecast accuracy in median *AFE* by 0.0100 relative to the ROE model, representing 31.4 percent of the total improvement of analyst consensus forecasts over the ROE model (reduction in median $AFE = 0.318$). In contrast, the corresponding figures for the OPINC and the CASHFLOW models are only 5.0 percent and 4.7 percent, respectively.

Taken together, these results in Table 5 indicate that earnings decomposition that recognizes cost variability and cost stickiness provides substantial incremental predictive content in forecasting one-year-ahead ROE relative to earnings decomposition that ignores such properties of earnings components. The evidence on the forecasting performance of our CVCS model relative to the more accurate analysts' consensus forecasts suggests a promising line of inquiry to examine if cost behavior information is a source of financial analysts' advantage over time-series models.

Loss Firms

Prior work in the accounting literature demonstrates that losses have lower earnings response coefficient (ERC) than profits, suggesting that the former is not as persistent as the latter due to the liquidation option (Hayn 1995). Our CVCS model contains a term that captures the differential impact of sales increase and decrease on future earnings. To the extent that the direction of sales changes is correlated with the sign of earnings, the observed superior forecasting performance of our CVCS model may be confounded.[23] To test this possibility, we augment the simple baseline ROE model (Equation (8)) to allow for differential persistence of profits and losses. This asymmetric ROE model is specified as follows:

$$ROE_t = \gamma_{e0} + \gamma_{e1}LOSS + \gamma_{e2}ROE_{t-1} + \gamma_{e3}LOSS{*}ROE_{t-1} + \varepsilon_t \qquad (8')$$

where $LOSS = \begin{cases} 1, & \text{if } ROE_{t-1} < 0 \\ 0, & \text{if } ROE_{t-1} \geq 0 \end{cases}$

We expect the revised ROE model in Equation (8′) to produce more accurate earnings forecasts than the simple ROE model in Equation (8) that does not distinguish between positive and negative earnings. More importantly, we expect the CVCS model to perform better than the asymmetric ROE model if our CVCS model captures an aspect of the earnings process beyond the liquidation option.

In estimating the asymmetric ROE model in Equation (8′) and forecasting one-year-ahead ROE, we follow the procedure described in the section titled "Forecast Model Estimation" to provide earnings forecasts on a comparable basis. Descriptive statistics on absolute forecast errors from the asymmetric ROE model are reported in Panel A of Tables 4 and 5. As expected, the asymmetric ROE model generates more accurate earnings forecasts than the simple ROE model. In Panel A, Table 4, for the full sample, the median (mean) *AFE* of the asymmetric ROE model is 0.0791 (0.1423), while the median (mean) *AFE* of the simple ROE model is 0.0805 (0.1427). Results reported in Panel A of Table 5 for the subsample of observations with available I/B/E/S forecasts are quantitatively similar to those in Table 4: the median (mean) *AFE* of the asymmetric ROE model is 0.0747

---

[23] We are grateful to an anonymous referee for suggesting we examine the possibly confounding effect of differential persistence of profits and losses.

(0.1202), while the median (mean) *AFE* of the simple ROE model is 0.0775 (0.1215). Such improvement in forecasting accuracy, however, does not fully explain the superior performance of our CVCS model: the median (mean) *AFE* of the CVCS model is 0.0734 (0.1412) for the full sample and 0.0675 (0.1168) for the I/B/E/S subsample. These results suggest that by incorporating asymmetric cost behavior, our CVCS model captures an aspect of the earnings process that is not a mere manifestation of the liquidation option.

V. ASSOCIATION WITH ABNORMAL STOCK RETURNS

In the previous two sections, we focus on forecast accuracy when comparing our CVCS model with the ROE, OPINC, and CASHFLOW models. Another performance criterion widely used in capital market research when evaluating earnings forecast models is based on the association of earnings forecast errors (i.e., unexpected earnings conditioned on a forecast model) with contemporaneous stock returns (Fried and Givoly 1982; Brown et al. 1987b; O'Brien 1988). This performance criterion reflects the informativeness (or information content) of an earnings forecast model in capturing investors' earnings expectations. Under our maintained assumption that the market's expectation is informationally efficient, the closer a selected model comes to mimicking the market's expectations, the less noise there is in forecasts from the model in representing unobservable market expectations, and thus the higher the association we expect to observe between the forecast errors from the model and the contemporaneous abnormal stock returns (Foster 1977). Therefore, the strength of the association between forecast errors and abnormal returns provides a yardstick to measure the relative information content of alternative earnings forecast models.

Prior research finds that forecast models with superior accuracy are not necessarily superior on the association dimension (Hughes and Ricks 1987; O'Brien 1988) so we evaluate the performance of the CVCS model based on the association between its forecast errors and contemporaneous stock returns, and compare it with the associations using the ROE, OPINC, CASHFLOW, and analyst models. We follow Dechow (1994) in evaluating the relative information content by using the likelihood ratio test proposed by Vuong (1989) for the following five simple regression models:

$$RET_t = \alpha_{a0} + \alpha_{a1} FE_t^{ROE} + \varepsilon_{at} \tag{20}$$

$$RET_t = \alpha_{b0} + \alpha_{b1} FE_t^{OPINC} + \varepsilon_{bt} \tag{21}$$

$$RET_t = \alpha_{c0} + \alpha_{c1} FE_t^{CASHFLOW} + \varepsilon_{ct} \tag{22}$$

$$RET_t = \alpha_{d0} + \alpha_{d1} FE_t^{CVCS} + \varepsilon_{dt} \tag{23}$$

$$RET_t = \alpha_{e0} + \alpha_{e1} FE_t^{IBES} + \varepsilon_{et} \tag{24}$$

where RET_t is the size- and book-to-market-adjusted contemporaneous annual stock returns (Fama and French 1992), and FE_t^{ROE}, FE_t^{OPINC}, $FE_t^{CASHFLOW}$, FE_t^{CVCS}, and FE_t^{IBES} are the forecast errors based on the ROE, the OPINC, the CASHFLOW, the CVCS models and I/B/E/S analyst consensus forecasts, respectively. The advantage of using Vuong's (1989) test is that it provides directional inferences as to which of two competing models is closer to the true underlying model without assuming that one model is the true one. The Vuong test is particularly suitable when the models being compared have significant explanatory power, as is true in our case (Vuong 1989; Dechow 1994).

We obtain information on the portfolio returns formed conditioned on size (market value of equity) and the book-to-market ratio to compute abnormal returns. The sample we use for the market association analysis is the intersection of the sample firms used above to compare earnings forecasts, and the CRSP database from which we obtain data on stock returns. We omit firms without December 31 fiscal year-ends and firm-year observations whose absolute abnormal stock returns exceed 100 percent. We also exclude observations with the most extreme 1 percent of forecast errors of each model. Our final sample is 8,745 firm-years from 2,747 firms from 1992 to 2002.[24]

Table 6 presents the correlation coefficients between forecast errors and abnormal stock returns. The Pearson (Spearman) correlations are above (below) the diagonal. The CVCS model produces forecast errors that are the most highly correlated with abnormal returns among the four time-series models we consider (Pearson $\rho = 0.350$, Spearman $\rho = 0.375$), while I/B/E/S analyst forecast errors are more highly correlated with abnormal returns (Pearson $\rho = 0.387$, Spearman $\rho = 0.459$) than any of the time-series model. The fact that the four time-series models generate highly correlated forecast errors is not surprising since they are conditioned on similar information sets, but the superior performance of the CVCS model relative to the other three time-series models obtains despite the near-unity correlations among them.

Table 7 contains results of the test of relative information content. Panel A reports regression results on the association between abnormal stock returns and forecast errors

TABLE 6
Correlation Coefficients[a] (p-values in parentheses) between Abnormal Stock Returns (RET)[b] and Forecast Errors (FE)[c] in Returns on Equity (ROE) Based on Four Time-Series Forecast Models and I/B/E/S Consensus Forecasts
Sample Period: 1992–2002 (n = 8,745)

	RET	FE^{ROE}	FE^{OPINC}	$FE^{CASHFLOW}$	FE^{CVCS}	FE^{IBES}
RET		0.342 (0.000)	0.338 (0.000)	0.331 (0.000)	0.350 (0.000)	0.387 (0.000)
FE^{ROE}	0.362 (0.000)		0.993 (0.000)	0.978 (0.000)	0.992 (0.000)	0.702 (0.000)
FE^{OPINC}	0.360 (0.000)	0.990 (0.000)		0.975 (0.000)	0.985 (0.000)	0.701 (0.000)
$FE^{CASHFLOW}$	0.348 (0.000)	0.971 (0.000)	0.965 (0.000)		0.973 (0.000)	0.697 (0.000)
FE^{CVCS}	0.375 (0.000)	0.984 (0.000)	0.975 (0.000)	0.959 (0.000)		0.723 (0.000)
FE^{IBES}	0.459 (0.000)	0.652 (0.000)	0.652 (0.000)	0.641 (0.000)	0.675 (0.000)	

[a] Pearson correlation coefficients are above the diagonal, and Spearman correlation coefficients are below the diagonal.

[b] Abnormal stock returns are measured as size and book-to-market adjusted returns, or annual raw returns less portfolio returns matched on size (market value of equity) and book-to-market (the ratio of book value of equity to market value of equity).

[c] FE^{ROE}, FE^{OPINC}, $FE^{CASHFLOW}$, FE^{CVCS}, and FE^{IBES} are forecast errors based on the ROE, OPINC, CASHFLOW, CVCS models, and I/B/E/S consensus analysts' forecasts, respectively.

---

[24] Results are similar to those reported here for the relative performance of the four time-series models when we use a larger sample of 17,170 firm-years that includes observations for which I/B/E/S data are not available.

TABLE 7

Regression Results of Abnormal Returns[a] on Forecast Errors (FE)[b] and Results of Vuong's Test of Relative Information Content of Four Time-Series Forecast Models and I/B/E/S Consensus Forecasts

Sample Period: 1992–2002 (n = 8,745)

Panel A: Simple Regression of Abnormal Returns on Forecast Errors (FE)

Independent Variable	Coeff. Est. (t-statistic)	Coeff. Est. (t-statistic)	Coeff. Est. (t-statistic)	Coeff. Est. (t-statistic)	Coeff. Est. (t-statistic)
Intercept	−0.002	−0.001	−0.001	0.013	0.025
	(−0.60)	(−0.26)	(−0.18)	(1.22)	(4.71)
FE^{ROE}	0.829				
	(34.03)				
FE^{OPINC}		0.818			
		(33.53)			
$FE^{CASHFLOW}$			0.803		
			(32.76)		
FE^{CVCS}				0.845	
				(34.95)	
FE^{IBES}					1.020
					(39.21)
Adj. R^2	0.1169	0.1139	0.1093	0.1225	0.1495

Panel B: Results of Vuong's Test of Relative Information Content

Competing Forecast Models		Vuong's Z-statistic[c]	p-value
ROE	vs. OPINC	−3.22	0.002
ROE	vs. CASHFLOW	−4.66	0.000
ROE	vs. CVCS	5.45	0.000
ROE	vs. I/B/E/S	4.76	0.000
OPINC	vs. CASHFLOW	−2.62	0.009
OPINC	vs. CVCS	6.29	0.000
OPINC	vs. I/B/E/S	5.22	0.000
CASHFLOW	vs. CVCS	7.12	0.000
CASHFLOW	vs. I/B/E/S	5.85	0.000
CVCS	vs. I/B/E/S	4.04	0.000

---

[a] Size and book-to-market adjusted returns equal annual raw returns less portfolio returns matched on size (market value of equity) and book-to-market (the ratio of book value of equity to market value of equity).

[b] FE^{ROE}, FE^{OPINC}, $FE^{CASHFLOW}$, FE^{CVCS}, and FE^{IBES} are forecast errors based on the ROE, OPINC, CASHFLOW, CVCS models, and I/B/E/S consensus analysts' forecasts, respectively.

[c] Vuong's Z-statistic is based on the likelihood ratio test developed by Vuong (1989) for non-nested model selection; a significant positive (negative) Z-statistic indicates that the first (second) forecast model is rejected in favor of the other forecast model.

in return on equity based on various forecast models. The highest adjusted R^2 of 0.1495 is achieved when abnormal returns are regressed on I/B/E/S analyst forecast errors, followed by forecast errors based on the CVCS model (adjusted R^2 = 0.1225). In Panel B, results based on pair-wise comparison indicate that the relative information content of the I/B/E/S analyst forecast is significantly greater than that of any of the mechanical forecast models. However, among the four time-series models we consider, Vuong's test rejects all three benchmark models when each of them is compared against the CVCS model. Specifically, Vuong's Z-statistic is 5.45 ($p < 0.001$) for the comparison between the ROE model and the CVCS model, 6.29 ($p < 0.001$) for the comparison between the OPINC model and the CVCS model, and 7.12 ($p < 0.001$) for the comparison between the CASHFLOW model and the CVCS model, all in favor of the CVCS model. In sum, these findings are consistent with the CVCS model having greater relative information content in measuring market's earnings expectations than other time-series forecast models that do not exploit the information in the relation between sales and cost items in the financial statements.

VI. CONCLUDING REMARKS

We propose an earnings forecast model (CVCS) that recognizes cost variability with sales changes and cost stickiness when sales decline. Unlike models of disaggregated earnings based on reporting classifications, our model stems from a management accounting perspective that emphasizes interactions among earnings components. The core of our model is the relation between sales and expenses embodied in the traditional fixed and variable cost behavior model. In addition, CVCS incorporates recent research that documents sticky cost behavior—the fact that costs do not decrease as much with sales decreases as they increase with sales increases.

Our CVCS model predicts one-year-ahead returns on equity better than do other models based on line items reported in income statements and statements of cash flows. Our CVCS model bridges over 30 percent of the information advantage financial analysts possess in forecasting one-year-ahead ROE over a basic time-series model that uses information on historical aggregate earnings. In contrast, the OPINC and CASHFLOW models bridge less than 6 percent of analysts' superior forecast accuracy over the basic ROE model. Invoking the assumption of market efficiency, we find that our model performs better than the other three time-series models when evaluated based on the association between the earnings surprise measured by forecast errors relative to these models and contemporaneous abnormal stock returns. Thus, the evidence suggests that earnings forecasts using our CVCS model better represent the market's earnings expectations than those using other models based on financial statement line items.

Our study raises several interesting questions to be explored in future research. As our association analysis is predicated on the maintained assumption of market efficiency, we do not address the issue of whether the stock market fully captures the information content of cost variability and cost stickiness relevant for forecasting future earnings. Although cost variability is a concept that is generally well understood, evidence on cost stickiness was largely anecdotal until recently. The extent to which investors have rational expectations about the impact of cost stickiness on future earnings over a longer horizon remains an open question.

Our objective is to consider a parsimonious model based only on income statement data. Promising research opportunities include expanding the scope of our model to incorporate additional information such as new product introduction or plant closures available to investors that may impact the earnings forecasts. In addition, the manner in which we model the impact of sticky cost behavior could be extended to consider its impact on the amount of accruals and related balance sheet items in order to develop a better understanding of the valuation process.

Our study documents that the simple cost variability and cost stickiness model has predictive content for the analysis of future profitability. Since such cost behavior models underlie some management accounting systems for budgeting and variance analysis, this

P28879

documentation is important in validating the basic premise supporting their design and indicating the importance of incorporating sticky cost behavior in improving their design. By emphasizing the importance of understanding cost behavior in forecasting earnings, our study points to the rich potential for future research integrating heretofore disparate streams of work in managerial and financial accounting.

REFERENCES

Anderson, M., R. Banker, and S. Janakiraman. 2003. Are selling, general and administrative costs "sticky"? *Journal of Accounting Research* 41: 47–63.

Brown, L. D. 1993. Earnings forecasting research: Its implications for capital markets research. *International Journal of Forecasting* 9: 295–320.

———, R. L. Hagerman, P. A. Griffin, and M. E. Zmijewski. 1987a. Security analysts superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics* 9: 61–87.

———, ———, ———, and ———. 1987b. An evaluation of alternative proxies for the market's assessment of unexpected earnings. *Journal of Accounting and Economics* 9: 159–193.

Collins, D., and P. Hribar. 2000. Earnings-based and accrual-based market anomalies: One effect or two? *Journal of Accounting and Economics* 29: 101–123.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18: 3–42.

———, S. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting Economics* 25: 133–168.

Fairfield, P., R. Sweeney, and T. Yohn. 1996. Accounting classification and the predictive content of earnings. *The Accounting Review* 71 (July): 337–355.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47: 427–465.

Foster, G. 1977. Quarterly accounting data: Time-series properties and predictive-ability results. *The Accounting Review* 52: 1–21.

Frankel, R., and C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics* 25: 283–319.

Fried, D., and D. Givoly. 982. Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics* 4: 85–107.

Garrison, R., and E. Noreen. 2002. *Managerial Accounting*. 10th edition. New York, NY: McGraw-Hill/Irwin.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20: 125–153.

Hughes, J., and W. Ricks. 1987. Association between forecast errors and excess returns near to earnings announcements. *The Accounting Review* 62: 158–175.

Kothari, S. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–231.

Lehmann, E. L. 1975. *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco, CA: Holden-Day, Inc.

Lev, B., and R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (Autumn): 190–215.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

Noreen, E., and N. Soderstrom. 1997. The accuracy of proportional cost models: Evidence from hospital service departments. *Review of Accounting Studies* 2: 89–114.

O'Brien. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10: 53–83.

Ou, J., and S. Penman. 1989. Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics* 11: 295–329.

Sheskin, D. 1997. *Handbook of Parametric and Nonparametric Statistical Procedures.* Boca Raton, FL: CRC Press.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71: 289–316.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57: 307–333.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 309–336

Pattern Identification and Industry-Specialist Auditors

Jacqueline S. Hammersley
University of Georgia

ABSTRACT: Complex financial statement misstatements that are difficult to diagnose are likely to be described by multiple cues that individually appear innocuous, but that together form an ominous pattern. Because individual auditors are likely to obtain only some of the cues forming the pattern, it is important to understand how well auditors interpret incomplete cue patterns. In this paper, I examine experimentally whether industry-specialist auditors develop problem representations about a seeded misstatement to facilitate interpretation of incomplete patterns. Overall, I find that matched specialists (i.e., those working in their industry) are able to interpret and fill in partial-cue patterns, whereas mismatched specialists do not recognize the implications of even full patterns. Matched specialists respond to a partial-cue pattern that potentially indicates misstatement by developing more complete problem representations about the seeded misstatement, assessing higher risk of misstatement, and suggesting procedures that will be efficient and effective at discriminating the presence of the seeded misstatement. Despite higher assessments of misstatement risk when receiving a partial-cue pattern, mismatched specialists' problem representations are impoverished and suggested procedures do not indicate a focus on the seeded misstatement.

Keywords: *auditor knowledge; industry specialization; pattern recognition; problem representations.*

Data Availability: *Data used in this study are available upon request.*

This paper is based on my dissertation, which was completed at the University of Illinois. I am very grateful to my dissertation committee members, Ira Solomon (Chair), Mark Peecher (Director of Research), Frederick Neumann, and Patrick Laughlin, for their guidance and direction. I am indebted to Gary Hecht for his able assistance in coding data. I also appreciate comments received from Michael Bamber, Jean Bedard, Jon Davis, Josh Herbold, Karla Johnstone, Kathryn Kadous, Natalia Kotchetova, Anne Magro, Bill Messier, Ken Trotman, Steve Salterio, Terry Warfield, two anonymous reviewers, workshop participants at the University of Georgia, Florida State University, Case Western Reserve University, University of Wisconsin–Madison, and University of Illinois, and conference participants at the 2004 Midyear Auditing Meeting, the 2004 AAA Annual Meeting, and the 2004 University of Illinois Symposium on Auditing Research. Finally, I am grateful to the auditors and their firms who generously gave their time to assist me.

Editor's note: This paper was accepted by S. Jane Kennedy, Editor and was the winner of the 2004 Competitive Manuscript Award.

Submitted January 2004
Accepted August 2005

I. INTRODUCTION

Auditors' failure to identify and report complex financial statement misstatements has had severe consequences for financial statement users and for auditors themselves. Financial statement misstatements that are complex[1] are hard to diagnose because they are likely to be described by a pattern of individually innocuous cues.[2] The problem of diagnosing such misstatements is further complicated when the pattern is incomplete. In this study, I examine whether industry-specialist auditors' unique knowledge structures allow them to diagnose such misstatements. Specifically, I investigate whether specialists use problem representations about the misstatement contained in incomplete industry-specific patterns to identify the increased risk present and to direct the search for evidence about the misstatement.[3]

It is important to understand the determinants of how well auditors interpret incomplete patterns suggestive of misstatement. Auditors must make judgments and decisions during the planning and evidence collection stages of the audit with incomplete information. Additionally, the division of labor among audit team members suggests the possibility that different auditors will collect the pieces of information that form a pattern. Individual auditors, who likely have seen only part of the pattern, would not have enough information, on their own, to identify the potential increased risk of misstatement. Consequently, unless at least one team member receives all of the information forming the pattern, the risk of misstatement may be assessed too low or identification of a misstatement may be delayed, negatively impacting both audit effectiveness and efficiency. Moreover, because individual pieces of information are not, in themselves, very suggestive of misstatement, they are unlikely to be communicated.[4]

I examine experimentally whether differences in industry specialists' problem representations about a seeded misstatement lead to task-performance differences. Industry-specialist auditors complete two cases, one matching and one not matching (as a benchmark) their industries of specialization. I manipulate whether they receive full, partial, or no patterns diagnostic of a misstatement embedded in case materials between participants. Auditors judge the likelihood of misstatement, explain the misstatement(s) about which they are concerned, determine necessary additional audit procedures, and perform surprise recalls for each case.

Consistent with expectations, I find that matched specialists develop more complete problem representations about the seeded misstatement when they receive partial- or full-cue patterns than when they receive no-cue patterns. Additionally, matched specialists who receive partial- and full-cue patterns assess the likelihood of misstatement higher than matched specialists who receive no-cue patterns. Critically, a mediation analysis confirms the role played by problem representations for matched specialists. Specifically, problem representation completeness significantly mediates the influence of the pattern manipulation on matched specialists' likelihood assessments. Matched specialists respond to a partial-cue pattern that potentially indicates misstatement by developing more complete problem representations about the seeded misstatement, assessing higher risk of misstatement, and suggesting procedures that will be efficient and effective at discriminating the presence of the seeded misstatement.

---

[1] Complexity denotes the number of alternatives or attributes that a decision maker has to consider. More alternatives or attributes are associated with more complex decisions (Payne et al. 1993).

[2] A pattern is composed of multiple pieces of information that must be combined, like the individual pieces of a mosaic, for the misstatement to emerge. When a pattern exists, no single piece of information is informative enough, on its own, to identify the misstatement. Rather, all of the cues forming the pattern must be perceived and appropriately processed for the misstatement to be conclusively identified.

[3] Problem representations are cognitive representations of complex phenomena that can be used to "run" thought experiments that allow inferences to be made (Christ 1993; Greeno 1989).

[4] The review process will result in formalized information sharing; however, evidence from court documents suggests that the review process does not always result in complete information sharing. For example, testimony from Lincoln Savings and Loan reveals that for the Continental Ranch land sale the manager in charge of the audit of AMCOR (the subsidiary recording the sale) was not aware that Lincoln had loaned the down payment to the buyer. This condition should have prohibited AMCOR from recording the $9 million gain on sale. Lincoln and AMCOR had separate audit teams; only the partners on the engagement team had responsibility for both entities (Erickson et al. 1996).

In contrast, mismatched specialists generally do not develop more complete problem representations about the seeded misstatement, even when they receive full-cue patterns. Additionally, despite assessing higher misstatement risk when receiving a partial-cue pattern versus no-cue pattern, mismatched specialists' suggested procedures do not focus on the seeded misstatement.

Collectively, these results suggest that industry specialists will improve audit quality when industry-specific knowledge is necessary to identify a misstatement and when knowledge of general accounting and audit testing, and thinking by analogy are not adequate substitutes for this knowledge. I document that one situation for which such industry-specialist knowledge is necessary is misstatement identification when a partial error pattern is present. To my knowledge, this is the first paper to investigate auditors' ability to identify partial patterns diagnostic of misstatement. This is a skill that is critical for successfully identifying complex financial statement misstatements, especially those being purposefully concealed.

The rest of this paper is organized as follows. The next section contains a review of the relevant literature and hypothesis development. In Section III, I describe the method that I use to test the hypotheses. In Section IV, I describe the experimental results. Finally, I discuss contributions and limitations.

II. LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Pattern Recognition

The ability to process patterns or configurations of stimuli that are important to subsequent judgments or decisions is known as configural information processing. Auditing researchers have reported that auditors configurally process information provided that all cues are present and domain-specific knowledge implies such a configuration is appropriate (Brown and Solomon 1990, 1991). Under these stylized conditions, auditors configurally process information about control-risk assessment, audit-procedure evidence evaluation, and financial statement account comparison analysis. Auditors also process information configurally when making estimates of extent of audit testing (Kerr and Ward 1994) and planning judgments (Maletta and Kida 1993). In addition, Bedard and Biggs (1991a) report that auditors generally attempt to use all the cues forming a pattern diagnostic of an overhead application error; however, they have difficulty generating correct hypotheses about the cause of the misstatement.

When patterns of information reveal misstatement, auditors risk issuing incorrect audit opinions unless they recognize the components of such patterns. When a multi cue pattern exists that is diagnostic of a material misstatement, auditors may receive none, some, or all of the cues. Previous research (e.g., Brown and Solomon 1990, 1991) that investigated auditors' ability to configurally process information manipulated the *level* of the cues received for a complete pattern (e.g., sales increased or sales decreased, a control was known to be present or absent). In contrast, I investigate how auditors respond to partial patterns of cues. If industry-specialist auditors possess superior knowledge about the industry in which they specialize and can use this knowledge to elaborate about the conditions under which missing information would be diagnostic of misstatement, they should be likely to identify a misstatement from a partial pattern.

Since auditors likely receive incomplete patterns, it is important that they can assess the related potential misstatement implications. To date, the literature does not address whether auditors are able to process patterns when the *presence or absence* of cues is manipulated (e.g., there is or is not information available about whether a specific control is in place to prevent a misstatement that could occur given the presence of another cue, and this information would be diagnostic about the risk of misstatement).

The question about this skill is especially important given the multi-person environment in which auditors work. This environment makes it likely that individual audit-team members receive partial-cue patterns. While the team collectively may receive all of the cues, no single team member may possess the full pattern. The team member receiving each cue must recognize its potential importance or communicate it to other team members in order to detect the misstatement. Alternatively, a single auditor ultimately may collect all the cues forming a pattern; identification of the pattern will be delayed until all the cues are collected if the auditor does not recognize the incomplete pattern.

Even if recognized, cues that comprise the pattern may go uncommunicated. Several features of the audit environment suggest that auditors are unlikely to communicate or share the information only they hold (i.e., unique information) with other team members, consistent with findings in psychology (Stasser and Titus 1985). First, auditors work in very high information-load conditions and such conditions reduce the likelihood that unique information is shared with other teammates (i.e., information sharing) (Stasser and Titus 1987). Second, larger auditees require larger audit teams and increases in group size reduce information sharing (Stasser et al. 1989). Finally, many audit tasks are "judgment" tasks for which a demonstrably correct answer does not exist (Kennedy et al. 1997). Stasser and Stewart (1992) report that when participants believe that their task has a demonstrably correct answer, they share more unique information than when they believe their task is a judgment task. Since conditions in the audit environment likely lead to low levels of unique information sharing, identification of partial patterns suggestive of misstatement is very important.

Industry-Specialists' Knowledge

Industry specialists are those individuals so designated by their firms and whose training and experience are mainly in a particular industry. Some recent research has investigated performance by industry specialists and auditors with high recent experience in a domain. For example, Owhoso et al. (2002) report that matched industry-specialist auditors detect more errors during the review process than do mismatched specialists. Taylor (2000) reports that banking industry-specialist auditors' inherent risk assessments differ (do not differ) from those of nonspecialists when evaluating loans receivable (property, plant, and equipment); he does not investigate whether these differences lead to differences in planned procedures. Bedard and Wright (1994) report that high-domain-experience auditors are more likely to react to current risk factors when planning audit procedures than are low-domain-experience auditors. Additionally, auditors with high-domain-experience generate more plausible hypotheses of likely errors than do auditors with low-domain-experience (Wright and Bedard 2000; Wright and Wright 1997).

However, matched specialists do not uniformly outperform mismatched specialists. Bedard and Biggs (1991b) report that high-domain-experience auditors who receive a correct management explanation are more likely than low-domain-experience auditors to recognize it as the correct hypothesis to explain a pattern. However, there is no difference between high- and low-domain-experience auditors when the management explanation is incomplete or missing, as is the case in the current study, suggesting that the benefits of specialization depend on the features of the situation. Finally, in the only study to explicitly investigate specialists' knowledge to date, Solomon et al. (1999) report that matched industry-specialist auditors know more non-errors and non-error frequencies than do mismatched industry specialists; however, results for error knowledge are mixed. This surprising result leads to concern that what industry specialists may learn is to explain away observed differences as non-errors, rather than recognize that they are errors and potentially affect audit effectiveness. This is consistent with the GAO's concern that industry specialists are too close to their clients and cannot objectively evaluate evidence (Carmichael 1997).

The prior literature suggests that industry experience is helpful in improving audit performance under some circumstances. I contend that use of industry specialists will improve audit quality when industry-specific knowledge is necessary to identify a misstatement and when knowledge of general accounting and audit testing, and thinking by analogy are not adequate substitutes for this knowledge. I contend that one situation for which such industry-specialist knowledge is necessary is misstatement identification when a partial error pattern is present. Specialists' cognitive representations may allow them to draw inferences from partial patterns of cues that are diagnostic of misstatement in their area of expertise.

Problem representations are cognitive representations of complex phenomena that people use to "run" thought experiments that allow inferences to be made (Greeno 1989). Problem representations are a function of external cues currently attended to and knowledge stored in memory (Brewer 1987). Problem representations develop naturally with experience in a domain and contain the causal features of the domain that they represent (Norman 1983; Rumelhart and Norman 1988). Domain knowledge plays an important role in the development of problem representations (Fincher-Kiefer et al. 1988; Spilich et al. 1979; Dutke 1986). Knowledge enables people to interpret information, allowing high-knowledge people to provide coherence and more extensive and accurate interpretations than low-knowledge people (Fincher-Kiefer et al. 1988; Voss et al. 1980). Additionally, high-knowledge people develop problem representations that are more elaborated that low-knowledge people (Tardieu et al. 1992).

In auditing, Christ (1993) investigated problem representation development in auditors of different ranks by examining free recalls. Her case did not contain a seeded misstatement or other manipulation of content that would cause an expected shift in focus of her participants. However, she reports that more experienced auditors' free recalls contain more abstractions (i.e., deeper principles) than less experienced auditors. I build on Christ's (1993) work in that I expect participants will focus on the seeded misstatements in my cases. I expect matched specialists to better interpret information about their industry than mismatched specialists when industry-specific knowledge, such as knowledge of unique accounting or reporting requirements or business practices, is required. Therefore, they should create better-developed problem representations focused on a seeded misstatement than do mismatched auditors. Bierstaker et al. (1999) report that problem representations that are focused on productive areas are critical to identifying seeded errors in analytical review.

Evidence about the existence and completeness of problem representations can be found in explanations and surprise recalls. People base their explanations, judgments, and recall of target events on their problem representation of the event. Problem representations contain the substance of the facts as well as inferences people make and believe to be part of the original event (Johnson-Laird 1983). Similarly, matched specialist auditors with knowledge of industry accounting practices and conditions will likely integrate case information related to a possible misstatement and consequently recall information focused on the seeded misstatement area. Additionally, matched specialists will elaborate about the implications of the information in the pattern and will recall more relationships between facts contained in the case and inferences about implications of the facts in the case (Christ 1993). Matched specialists' elaborations about the case will also result in explanations about what (if anything) could be wrong that are focused on the misstatement area and diagnose the seeded misstatement.

Therefore, I expect matched specialists who receive full- and partial-cue patterns to be able to understand the relationships among the cues. Recipients of partial-cue patterns are expected to be able to provide coherence to the incomplete pattern. This should lead partial-cue pattern recipients to elaborate about the incomplete pattern and hypothesize about the missing information. They may even infer the existence of the missing piece of the cue pattern. Matched specialists who receive none of the cue pattern are not expected to elaborate about the unstated error pattern. As a result, matched specialists who receive partial-cue patterns are expected to have better-developed problem representations about the seeded misstatement than will matched specialists who receive no-cue patterns.

Matched specialists who receive complete cue patterns are also expected to elaborate about the cue pattern and the error it suggests. As a result, matched specialists who receive complete cue patterns will have better developed problem representations about the seeded misstatement than will those who receive no-cue patterns; matched specialists who receive full-cue patterns will have problem representations that are at least as well developed as matched specialists who receive partial-cue patterns. Figure 1 depicts my predictions about problem representation development. More formally, I predict (all hypotheses are stated in alternative form):

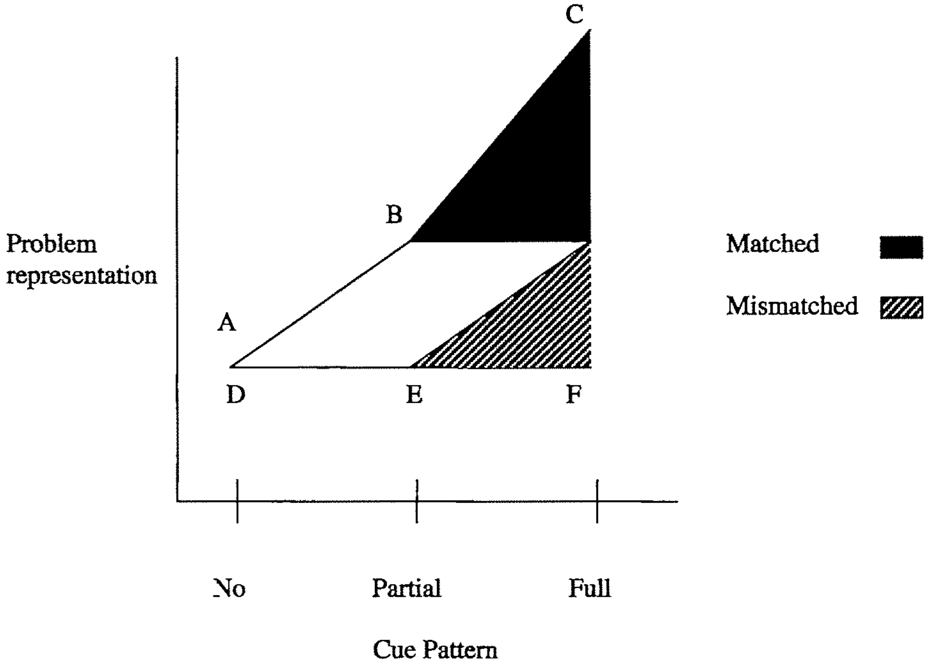
H1a: Matched specialists receiving partial- and full-cue patterns will have better developed problem representations about the seeded misstatement than will matched specialists receiving no-cue patterns.

Mismatched specialists who receive partial-cue patterns are not expected to be able to fully understand the implications of the incomplete information. They are not expected to elaborate on the incomplete pattern because they are not expected to recognize its significance. Their problem representations are not expected to differ in level of development from mismatched specialists who receive no-cue patterns. Mismatched specialists who receive full-cue patterns may be able to interpret the patterns because they may have enough information, combined with their general accounting knowledge, to deduce the misstatement. Consequently, if they recognize the significance of the pattern, then they are expected to develop problem representations that are better developed than mismatched specialists who receive the partial-cue pattern. If they are unable to recognize the significance of the pattern because only industry-specific knowledge is sufficient to identify the pattern, then their problem representations will be at least as well developed as those of matched specialists who receive partial-cue patterns. More formally, I predict:

H1b: Mismatched specialists receiving full-cue patterns will have problem representations about the seeded misstatement that are at least as well developed as those of mismatched specialists who receive no- or partial-cue-patterns.

FIGURE 1
Hypothesis 1

H1a contrasts: $A < (B + C)/2$
H1b contrasts: $(D + E)/2 \leq F$
H1c contrasts: $B > E$



Industry match is operationalized at two levels within participants. Matched participants are those working in their industry of specialization; mismatched participants are those working outside their industry of specialization.
The *cue-pattern* condition is operationalized at three levels between participants. Participants in the None condition received no cues from the seeded misstatement pattern. Participants in the Partial (Full) condition received two (all three) cues from the seeded misstatement pattern.
The *dependent variable* is the sum of four standardized measures of the completeness of the participants' problem representations about the seeded misstatement. A higher number indicates better developed problem representations.

The above discussion suggests that the key difference between matched and mismatched specialists will occur when they receive partial-cue patterns. Matched specialists who receive partial-cue patterns are expected to understand the relationship among the cues and provide coherence to incomplete patterns resulting in well-developed problem representations about the seeded misstatement. Mismatched specialists who receive partial-cue patterns are not expected to understand these relationships, resulting in relatively impoverished problem representations about the seeded misstatement. More formally, I predict:

H1c: Matched specialists receiving partial-cue patterns will have better developed problem representations about the seeded misstatement than will mismatched specialists receiving partial-cue patterns.

Likelihood Assessments

Matched specialists likely possess knowledge of the patterns of business-operations information that would be diagnostic of misstatement, such as departures from expectations about "normal" conditions. Knowledge of these patterns may be acquired directly by experiencing the pattern on audits or indirectly from training sessions or discussions with colleagues. Auditors with this knowledge would have the cues that comprise the pattern accessibly stored in memory with connections developed between the cues that compose the pattern. Therefore, even if specialists receive only some of the cues comprising the pattern, it is likely that they would recognize the elevated potential for misstatement. Additionally, auditors with well-developed problem representations likely will elaborate about the conditions under which the information heightens the risk of material misstatement, even if they have not previously encountered the misstatement.

Matched specialists who receive partial- and full-cue patterns suggestive of misstatement will assess the likelihood of misstatement conditioned on two factors: the perceived diagnosticity of the partial-cue pattern about error and the conditional likelihood of the error(s) they are concerned about relative to other possible non-error conditions given the partial pattern. Matched specialists who receive no-cue patterns suggestive of misstatement are not expected to perceive the unstated error risk. Therefore, I expect that matched specialists who receive partial-cue patterns will assess the likelihood of misstatement higher than matched specialists who receive no-cue patterns.

Matched specialists who receive full-cue patterns will assess the likelihood of misstatement higher than those who receive no-cue patterns; however, their likelihood assessments relative to those in the partial-cue pattern condition will depend on the relative diagnosticity of the partial- and full-cue patterns. When the full pattern is much more diagnostic of misstatement than is the partial pattern, those receiving the full pattern will produce likelihood assessments that are higher than those receiving partial patterns. When the full pattern is only marginally more diagnostic of misstatement than is the partial pattern, those receiving full patterns will assess the likelihood of misstatement at least as high as those who receive partial-cue patterns. Overall I expect auditors' likelihood assessments to be of the form depicted in Figure 2 and as stated more formally below.

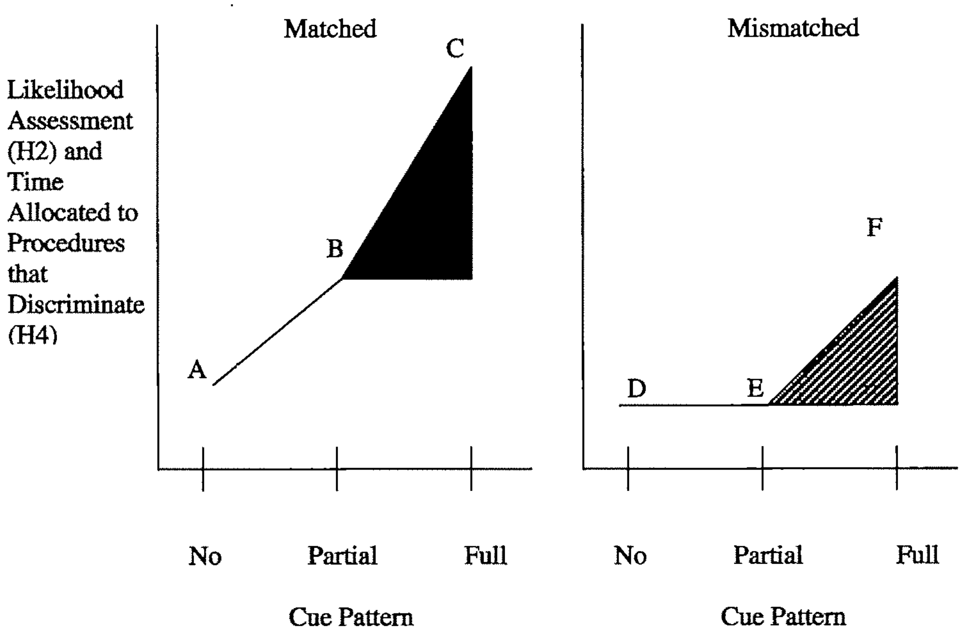
H2a: Matched specialists receiving partial- and full-cue patterns will assess higher likelihoods of misstatement than will matched specialists receiving no-cue patterns.

Mismatched specialists who receive partial-cue patterns are also expected to assess the likelihood of misstatement conditioned on the perceived diagnosticity of the partial-cue pattern about error and the conditional likelihood of the error(s) they are concerned about relative to other possible non-error conditions given the partial pattern. However, because they are expected to develop relatively impoverished problem representations, mismatched specialists who receive partial-cue patterns are not likely to perceive the diagnosticity of the partial-cue pattern and, thus, are not expected to increase their assessment of risk of material misstatement. Their likelihood assessments are not expected to differ from those of mismatched specialists who receive no-cue patterns. Mismatched specialists who receive full-cue patterns may perceive the diagnosticity of the pattern for misstatement and if they do, they are expected to increase their assessed likelihood of misstatement accordingly. If they are not able to recognize the diagnosticity of the pattern for misstatement, their likelihood assessments are not expected to differ from those of mismatched specialists who receive partial-cue patterns. More formally, for mismatched specialists I predict

---

FIGURE 2
Hypotheses 2 and 4

H2a and H4a contrasts: $A < (B + C)/2$
H2b and H4b contrasts: $(D + E)/2 \leq F$


Likelihood Assessment (H2) and Time Allocated to Procedures that Discriminate (H4)
Matched
C
B
A
Mismatched
F
D
E
No
Partial
Full
No
Partial
Full
Cue Pattern
Cue Pattern

See Figure 1 for explanations of industry match and cue-pattern variables.
The dependent variable for H2 is the likelihood that a material misstatement was contained in the financial statements. Participants rate this likelihood on a 101-point scale anchored by 0 (No chance of material misstatement) and 100 (Certain of material misstatement). I present two graphs as theory does not specify the relative positions of the lines for matched and mismatched participants for this dependent variable.
The dependent variable for H4 is the total time allocated to procedures coded as discriminatory. Participants listed procedures they believed were necessary to determine whether the potential misstatement(s) they had listed existed and time necessary to complete those procedures.

H2b: Mismatched specialists receiving full-cue patterns will assess the likelihood of misstatement at least as high as mismatched specialists receiving no- or partial-cue patterns.

The preceding discussion specifies predictions about likelihood assessments of mismatched specialists relative to other mismatched specialists. I make no predictions about likelihood assessments of matched versus mismatched specialists as this comparison is not clear *ex ante*. For example, one response to the uncertainty associated with being mismatched may be to assess the likelihood of misstatement higher overall.

Above I examined the predicted impact of pattern completeness on likelihood assessments. The discussion hypothesized that pattern completeness will affect likelihood assessments in an indirect manner through the effect of pattern completeness on specialists' problem representation development. I expect that problem representation development will directly affect specialists' likelihood assessments. More formally, I predict:[5]

> **H3:** Matched specialists' problem representations about the seeded misstatement are expected to mediate the relationship between the pattern completeness manipulation and likelihood assessments.

Audit Procedures

As discussed earlier, the receipt of cues that are diagnostic of misstatement is expected to affect industry specialists' assessed likelihood of misstatement. Industry specialists who suspect a misstatement pattern likely seek information that can be used to determine whether the full pattern exists and the extent of any misstatement. Specialists who believe that there is a relatively higher risk of misstatement are expected to know what industry-specific conditions would give rise to a misstatement, what tests should be performed, and what evidence should be gathered to determine the existence and extent of misstatement.

More experienced auditors perform more directed searches for information and choose procedures that yield evidence that better discriminates among possible causes of unexpected differences than do less experienced auditors (Biggs et al. 1988). Similarly, I expect specialists who are more concerned about a particular misstatement to propose a more directed search for evidence about that misstatement (i.e., to seek information with a greater ability to discriminate between the hypothesized error and a non-error). More audit time, therefore, should be allocated to evidence with higher discriminatory power.

Matched specialists who receive partial-cue patterns will elaborate about the incomplete pattern, hypothesize about the missing information, and conduct more directed searches for evidence than will matched specialists who receive no-cue patterns. Consequently, I expect these specialists to allocate more time to procedures that will discriminate whether the seeded misstatement is present than will matched specialists who receive no-cue patterns.

Matched specialists who receive complete cue patterns are also expected to elaborate about the cue pattern and the error it suggests. As a result, matched specialists who receive complete cue patterns will allocate more time to procedures that will discriminate whether the seeded misstatement is present than those who receive no-cue patterns; however, the amount of time allocated relative to those in the partial-cue pattern condition will depend on the relative diagnosticity of the partial- and full-cue patterns. When the full pattern is much more diagnostic of misstatement than is the partial pattern, those receiving full patterns will allocate more time to discriminatory procedures than those receiving partial patterns. However, when the full pattern is only marginally more diagnostic of misstatement than is the partial pattern, those receiving full patterns will allocate at least as much time to discriminatory procedures as will those receiving partial patterns. Figure 2 depicts my predictions about time allocated to discriminatory procedures. More formally, I predict:

> **H4a:** Matched specialists receiving full- and partial-cue patterns will allocate more time to procedures that will discriminate whether the seeded misstatement is present than will matched specialists receiving no-cue patterns.

---

[5] I do not make a similar mediation prediction for mismatched specialists as I do not expect there will be sufficient variation in mismatched specialists' problem representation development or likelihood assessments to allow detection of such a relationship.

Mismatched specialists who receive partial-cue patterns are not expected to be able to interpret the pattern, nor are they expected to understand the industry-specific conditions that would give rise to the misstatement. As such, they are not expected to have a particular misstatement in mind about which they will seek evidence. Therefore, mismatched specialists who receive partial-cue patterns are not expected to differentially allocate time to procedures that will discriminate whether the target misstatement is present compared to mismatched specialists who receive no-cue patterns. Mismatched specialists who receive full-cue patterns may be able to identify the potential misstatement and may be able to conduct directed searches for evidence if their general knowledge of audit testing allows them to design tests that would discriminate the misstatement. If only industry-specific knowledge is sufficient to design tests that will discriminate the presence of the seeded misstatement, then they will allocate at least as much time to discriminatory procedures as mismatched specialists who receive partial-cue patterns. Stated more formally, these predictions are:

H4b: Mismatched specialists receiving full-cue patterns will allocate at least as much time to procedures that will discriminate whether the seeded misstatement is present as will mismatched specialists receiving no- or partial-cue patterns.

Specialists who believe that there is a relatively low risk of misstatement are unlikely to have specific misstatements in mind. Consequently, they will be relatively unlikely to identify audit procedures that would provide evidence about the presence of a particular misstatement. Less experienced auditors generally perform ill-defined searches and choose audit procedures that poorly discriminate among possible causes of audit fluctuations (Biggs et al. 1988). Consequently, I expect mismatched auditors to allocate more time to audit procedures with relatively low ability to discriminate whether a misstatement is present than do matched specialists. Stated more formally, this prediction is:

H5: Mismatched specialists will allocate more time to procedures unlikely to discriminate the presence of misstatement than will matched specialists.

III. RESEARCH METHOD

Participants

Participants volunteered after being identified as firm-designated industry-specialist auditors in banking or state and local governments. Seventy-four industry-specialist auditors from four Big 5 firms participated in one of 19 sessions of the experiment;[6] responses from 65 auditors are reported herein.[7] As summarized in Table 1, participants' mean audit experience is 83.82 months. Participants are partners/principals (n = 10), senior managers (n

---

[6] I collected data from 44 participants pre-Enron (35 banking specialists and 9 government specialists) and 21 government specialists post-Enron. Because of the potential effect on the data I performed careful analysis to ensure that there were no systematic differences across time periods. I estimated all of the models including timing of data collection (pre- or post-Enron) as a factor nested within specialty. The timing of data collection was not significant in any of the models for any dependent variable and the timing did not interact with the other independent variables in the models. Additionally, I re-estimated the contrasts used to test the hypotheses and all statistical inferences hold when timing of data collection is included in the models. Consequently, I conclude that timing of data collection did not impact the dependent variables and it is appropriate to exclude timing from the models and report the more parsimonious models.

[7] I exclude nine participants because they reported either significant experience in both industries (n = 2) or no significant experience in either industry (n = 7).

TABLE 1
Participant Demographics

	All Participants (n = 65)	Banking Specialists (n = 35)	Government Specialists (n = 30)
Months of audit experience [Mean (Std. Dev.)]	83.82 (69.87)	71.46 (62.40)	98.23 (76.22)
Title			
Seniors	30	19	11
Managers	14	9	5
Senior Managers	11	2	9
Partners/Principals	10	5	5
Months experience auditing			
Banks		49.09 (52.75)	4.87 (15.62)
State and local governments		0.57 (2.16)	72.67 (78.94)
Percent of time in past year auditing			
Banks		63.46% (29.79)	0.03% (0.18)
State and local governments		0.43% (2.54)	56.60% (29.63)

= 11), managers (n = 14), and seniors (n = 30).[8] Banking (government) specialists report having an average of 49.09 (4.87) months of experience auditing banking clients and 0.57 (72.67) months of experience auditing government clients. Banking (government) specialists report spending an average of 63.46 percent (0.03 percent) of their time in the past year auditing banking clients and 0.43 percent (56.60 percent) of their time in the past year auditing government clients.

Task

The experimental instrument contains two cases, one bank and one city-government case, developed in consultation with audit partners who specialize in each industry. I chose these industries to minimize expected knowledge transfer, thus providing a strong test of the effects of industry-knowledge differences on performance. These industries have unique accounting issues and the industries are not included in the same industry category by any of the Big 5 accounting firms. For each case participants were asked to: read the case, assess the likelihood of material misstatement, report what (if any) misstatement they were concerned about, list procedures they would perform and audit hours necessary to determine whether their noted misstatements were present, and recall all important information from each case. After completing the second case, participants answered a post-experimental questionnaire that contained manipulation checks and demographic questions.

---

[8] The participating firms designate qualifying auditors as industry specialists beginning at the senior level. I reran all the analysis for H1, H2, H4, and H5 including total months of audit experience as a covariate. Total months of audit experience was not significant in any of the models as a covariate and it did not interact with any of the other variables in the models. Additionally, all inferences hold when the variable is included. This analysis provides evidence that the results are not driven by the more experienced participants but can be attributed to auditor specialization.

Participants received unaudited revenue numbers and background information about two revenue items that comprised the area under audit in each case. The cases are presented in Appendices A and B. Each case contained a seeded material misstatement of one of the two revenue items discussed. The revenue recognition problem of interest was unique to, but common within, the target industry. This ensured that industry knowledge, rather than general business or accounting knowledge, was necessary to identify the risk of misstatement.[9] A series of manipulation-check questions confirm that matched specialists knew the appropriate accounting for the different types of revenue and that mismatched specialists were not able to reason by analogy to determine the correct accounting for the target revenue areas. In the post-experimental questionnaire, participants were asked to recognize the appropriate accounting method for each of the five types of revenue that appear in the cases (i.e., successful loan application fees, unsuccessful loan application fees, credit card fees, grant revenue, and sewer taxes). Participants answered on an 11-point scale anchored by 0 (certain does not comply with GAAP) and 10 (certain does comply with GAAP). The target revenues in the case were the successful loan application fees and the grant revenue. These items required knowledge of an industry-specific accounting method and of the characteristics that allow one to determine the correct application of the method.[10] I expected this knowledge would not easily be transferred from another domain or solved by analogy. The non-target revenue types, although still specific to the industry, did not require an industry-specific accounting method. I expected that participants, using their general accounting knowledge, would be able to reason by analogy and be confident that they knew how to account for the non-target revenues.

The data support these expectations. Government specialists rated the appropriate accounting method for the grant revenue higher (7.80) than did the banking specialists (3.47) ($p = 0.00$); banking specialists rated the appropriate accounting method for the successful loan application-fee revenue higher (5.98) than did the government specialists (3.82) ($p = 0.01$). These differences in knowledge of the correct accounting for the target revenue items suggest that this knowledge was necessary to allow matched specialists to develop more complete problem representations about the target misstatements. Additionally, it provides evidence that mismatched specialists were not able to transfer knowledge or reason by analogy to determine the correct accounting treatments for the target revenue areas. There were no significant differences between the two groups of specialists on the other three revenue measures.[11] This suggests that mismatched specialists were able to transfer some knowledge or reason by analogy and apply it to these non-target revenues.

Participants received the experimental materials in a set of six numbered envelopes. They were instructed to work at their own pace, to return each completed document to its envelope, and to close it before opening the next envelope. A proctor ensured participants complied with instructions during all sessions of the experiment. On average, participants completed the materials in about 75 minutes.

---

[9] Matched specialists had previously encountered the seeded errors 16.9 percent of the time; mismatched specialists had encountered the seeded errors only 1.5 percent of the time. This difference is significant ($p < 0.01$), confirming that matched specialists had more experience with the seeded errors than did mismatched specialists.

[10] For the banking case, participants had to determine, from the characteristics given, whether the loan fees were direct; for the government case, participants had to determine, from the characteristics given, whether the grant was reimbursement driven.

[11] Government specialists did not rate the appropriate accounting method for the sewer taxes higher (6.06) than did the banking specialists (5.19) ($p = 0.32$). Banking specialists did not rate the appropriate accounting method for the unsuccessful loan application-fee revenue higher (8.71) than did the government specialists (8.05) ($p = 0.22$); banking specialists did not rate the appropriate accounting method for credit card fees higher (8.71) than did the government specialists (8.09) ($p = 0.26$).

Independent Variables

There are two independent variables: industry match and completeness of the misstatement cue pattern.[12] I create the two levels of industry match (i.e., matched and mismatched) by manipulating context and measuring participant industry specialization. I manipulate context within-participants at two levels. Each participant evaluates one banking case and one city-government case.[13] I measure participants' industry of specialization; participants specialize in auditing banks or state and local governments. I manipulate completeness of the misstatement cue pattern between-participants. The misstatement pattern for each case consists of three cues whose presence or absence in the case is manipulated at three levels (i.e., no, zero cues present; partial, two cues present; and full, all three cues present). The pattern cues must be combined with the case background information to determine whether a misstatement exists. The background information, on its own, provides no reason for participants to suspect the misstatement. The partial pattern for each case provides information that allows participants to start forming an expectation about the amount of revenue that should be recognized and to determine that under certain circumstances there is a particular misstatement. However, they must receive the information in the third cue to be certain of the misstatement. The third cue for each case contains information about a client-specific accounting convention. This information, combined with participant knowledge about how the transaction should be accounted for, allows participants to determine the correct misstatement.

Dependent Variables

Three classes of dependent variables were measured: problem representation measures, likelihood assessments, and total time allocated to procedures that will or will not discriminate the misstatement. I used four different problem representation measures to create the overall measures used to test H1 and H3: recall content, recall character, explanation content, and explanation direction. Higher scores are consistent with better-developed problem representations.

I computed the recall content and character measures from participants' listing of all the important information they could remember from the case. Participants were instructed not to look at the case materials while performing this task. I parsed the recalls into idea units, and a doctoral student with auditing experience and I worked independently to code the recalls and all other items for the project. The doctoral-student coder was blind to participant specialty and the hypotheses for all coding for this project.

We coded recall idea units on two dimensions: content and character. Content of each recall idea unit was coded as relating to the misstatement area or relating to other areas of the case. Items coded as related to the misstatement area were further coded as to their character; each idea unit was coded as expressing either case facts or relationships and inferences from the case. Participants who were better able to interpret the case scenario as indicating misstatement were expected to recall more items related to the misstatement area. These participants also were expected to make connections among facts presented and include more relationships and inferences related to the misstatement area in their recalls. Inter-rater agreement is 97.9 percent for recall content and 89.3 percent for character. Cohen's Kappa, a measure of inter-rater agreement over and above that expected by random agreement, is 0.969 and 0.626 for content and character, respectively; both are significantly different from zero ($p < 0.001$). All coding differences for the project were mutually resolved.

---

[12] I also manipulate the order of the cases; half of the subjects receive the banking case first and half receive the government case first. Except as noted there are no predicted or observed order effects, so I combine all subsequent analysis across order conditions.

[13] I also analyzed the results by case and the results are qualitatively the same as those reported in the paper. The predictions, in general, hold for each hypothesis for each case, however, there is reduced significance in some instances due to the reduction in sample size and resulting statistical power. I had no *ex ante* predictions or theoretical reason to investigate differences due to a specific industry, so I report the results combined over the two industries to take advantage of the experimental design and rule out certain alternative explanations for my results (e.g., that something idiosyncratic about one case or specialty is driving the average results, or that specialists in one industry have higher or lower ability and this drives the average result).

I computed the explanation content and direction measures from participants' explanations of the misstatement(s), if any, about which they were most concerned. We coded these responses on two dimensions: content and direction. The content of each explanation idea unit was coded as relating to the misstatement area or other areas of the case. The direction of each explanation idea unit was coded as supporting or not supporting the seeded misstatement about the target revenue area. Participants who were better able to interpret the case scenario and who focused on the seeded misstatement were expected to explain more items that relate to the target revenue area. The participants also were expected to include more items that explain the seeded misstatement. Inter-rater agreement is 90.6 percent for explanation content and 79.3 percent for direction. Cohen's Kappa is 0.857 and 0.683 for content and direction, respectively; both are significantly different from zero ($p < 0.001$).

After reading each case and before listing the recalls, participants rated the likelihood that a material misstatement was contained in the financial statements. The likelihood assessment is the dependent variable for H2. Participants were free to look back at the case materials while making this assessment. The likelihood assessment is elicited on a 101-point scale anchored by 0 (no chance of material misstatement) and 100 (certain of material misstatement).

Finally, participants listed procedures and specified the audit time necessary to determine whether the potential misstatements they had previously identified were present. We coded these procedures on the ability to discriminate the target misstatement. The total time allocated to discriminatory procedures is the dependent measure for H4; total time allocated to nondiscriminatory procedures is the dependent measure for H5. Procedures were coded as discriminatory if they would provide direct evidence about whether the target misstatement is present; procedures were coded as nondiscriminatory if they would not provide such evidence or if they would provide evidence about the non-target revenue area. Inter-rater agreement is 79.5 percent; Cohen's Kappa is 0.514 and is significantly different from zero ($p < 0.001$).

IV. RESULTS

Hypothesis 1

In H1a, I predict that matched participants' problem representations about the seeded misstatement will be better developed when they receive full- and partial-cue patterns than when they receive no-cue patterns. In H1b, I predict that mismatched participants' problem representations about the seeded misstatement will be at least as well developed when they receive full-cue patterns as when they receive no- or partial-cue patterns. Finally in H1c, I predict that matched specialists receiving partial-cue patterns will have better-developed problem representations about the seeded misstatement than will mismatched specialists receiving partial-cue patterns. The hypotheses predict specific patterns of cell means, so I use contrasts as the primary tests of the hypotheses (Buckless and Ravenscroft 1990).

I estimate a repeated-measures ANOVA with the overall problem representation measures as dependent variables and industry match (within participant) and cue pattern condition as independent variables. I present the means and the results of the planned contrasts in Table 2, Panels A and B, respectively. Matched participants who receive full- or partial-cue patterns have better developed problem representations for the seeded misstatement (mean 1.26) than do matched participants in the no-cue pattern condition (−0.51) and this difference is significant ($F_{1,62} = 7.545$, one-sided $p = 0.004$), supporting H1a. Additionally, a *post hoc* test reveals that the problem representation development of participants in the

TABLE 2
Problem Representation Measures
LS Mean

H1a contrast: A < (B + C)/2
H1b contrast: (D + E)/2 ≤ F
H1c contrast: B > E

Panel A: Problem Representation Measures

		Cue Pattern		
		No **n = 21**	**Partial** **n = 23**	**Full** **n = 21**
Matched	Recall Content	−0.09	0.13	0.44
	Recall Character	0.20	0.15	0.28
	Explanation Content	−0.44	0.01	0.30
	Explanation Direction	−0.17	0.59	0.62
	Overall Measure	−0.51	0.89	1.64
	(SE)	(0.53)	(0.51)	(0.53)
	Cell	A	B	C
Mismatched	Recall Content	−0.46	−0.19	0.25
	Recall Character	−0.14	−0.16	−0.22
	Explanation Content	−0.17	−0.09	−0.06
	Explanation Direction	−0.62	−0.37	−0.45
	Overall Measure	−1.39	−0.80	−0.48
	(SE)	(0.45)	(0.43)	(0.45)
	Cell	D	E	F

Panel B: Tests of H1

Planned Contrasts	$\mathbf{F_{1,62}}$	**p > F** **(one-tailed)**
Matched No < (Matched Partial + Full)/2	7.545	0.004
(Mismatched No + Partial)/2 ≤ Mismatched Full	1.293	0.130
Matched Partial > Mismatched Partial	7.856	0.004

Industry match is operationalized at two levels within participants. Matched participants are those working in their industry of specialization; mismatched participants are those working outside their industry of specialization.

The *cue-pattern* condition is operationalized at three levels between participants. Participants in the No Cue Pattern condition received no cues from the seeded misstatement pattern. Participants in the Partial- (Full-Cue) Pattern condition received two (all three) cues from the seeded misstatement pattern.

The *dependent variable* is computed by standardizing the four problem representation measures (Recall content, Recall character, Explanation content, and Explanation direction) using Z-scores and summing to compute overall problem representation measures.

partial-cue pattern condition (0.89) is greater than that of those in the no-cue pattern condition ($F_{1,62} = 3.622$, one-sided $p = 0.031$). Mismatched participants who receive full-cue patterns have problem representations for the seeded misstatement (−0.48) that are at least as well developed as those of mismatched participants in the no- or partial-cue pattern conditions (−1.10) ($F_{1,62} = 1.293$, one-sided $p = 0.130$), consistent with H1b. Finally, matched participants in the partial pattern condition have better developed problem representations for the seeded misstatement (0.89) than do mismatched participants in the partial pattern condition (−0.80) ($F_{1,62} = 7.856$, one-sided $p = 0.004$), supporting H1c.

Recall and Explanation Problem Representation Measures

I also investigate whether the overall problem representation results are consistent across its components by disaggregating the overall measure into recall and explanation problem representation measures. I compute a total recall problem representation measure by summing the recall content and character measures. I compute a total explanation problem representation measure by summing the explanation content and direction measures. I estimate a repeated-measures ANOVA with the total recall and total explanation problem representation measures as dependent variables and industry match (within participant) and cue pattern condition as independent variables. Matched participants who receive full- or partial-cue patterns have more developed recall problem representation measures related to the seeded misstatement (mean 0.50) than do matched participants in the no-cue pattern condition (0.11) but this difference is not significant ($F_{1,62} = 0.707$, one-sided $p = 0.202$). The pattern of means is consistent with the hypothesis for the recall content dimension of this measure. The insignificant result appears to be driven by the character dimension of this measure; matched specialists appear to generate inferences and relationships related to the misstatement area even in the no-cue pattern condition.

Matched participants who receive full- or partial-cue patterns have more developed explanation problem representation measures related to the seeded misstatement (mean 0.76) than do matched participants in the no-cue pattern condition (−0.61) and this difference is significant ($F_{1,62} = 10.431$, one-sided $p = 0.001$). Additionally, a *post hoc* test reveals that the explanation development of participants in the partial-cue pattern condition (0.60) is greater than that of those in the no-cue pattern condition ($F_{1,62} = 6.326$, one-sided $p = 0.008$). Overall, this evidence is consistent with H1a; matched specialists have better developed problem representations in the partial- and full-cue pattern conditions than in the no-cue pattern condition.

Mismatched participants who receive full-cue patterns have recall problem representations related to the seeded misstatement (0.03) that are at least as well developed as those of mismatched participants in the no- and partial-cue pattern conditions (mean −0.47) ($F_{1,62} = 1.692$, one-sided $p = 0.099$). Mismatched participants who receive full-cue patterns have explanation problem representations related to the seeded misstatement (−0.51) that are at least as well developed as those of mismatched participants in the no- and partial-cue pattern conditions (mean −0.63) ($F_{1,62} = 0.141$, one-sided $p = 0.354$). This evidence is consistent with H1b.

Finally, matched participants in the partial-cue pattern condition have better developed recall problem representations for the seeded misstatement (0.28) than do mismatched participants in the partial-cue pattern condition (−0.35) and this difference is significant at the 10 percent level ($F_{1,62} = 2.396$, one-sided $p = 0.063$). Matched participants in the partial-cue pattern condition have better developed explanation problem representations (0.60) than do mismatched participants in the partial-cue pattern condition (−0.46) and this difference is significant ($F_{1,62} = 6.931$, one-sided $p = 0.006$). This evidence supports H1c; matched specialists have better developed problem representations in the partial-cue pattern condition than do mismatched specialists.

Hypothesis 2

In H2a, I predict that matched participants will assess the likelihood of misstatement higher when they receive partial- or full-cue patterns than when they receive no-cue patterns. In H2b, I predict that mismatched participants' likelihood assessments will be at least as high when they receive full-cue patterns as when they receive no- or partial-cue patterns.

I estimate a repeated-measures ANOVA with matched and mismatched likelihood of misstatement as the dependent variables and industry match (within participant) and cue-pattern condition (between-participant) as independent variables. Table 3 provides descriptive statistics (Panel A) and planned contrast results (Panel B). Cell means are graphed in Figure 3. Matched participants' mean likelihood assessments in the partial- and full-cue pattern conditions (66.12) is greater than those in the no-cue pattern condition (43.90) and this difference is significant ($F_{1, 62} = 6.720$, one-sided $p = 0.006$). Additionally, a *post hoc* test reveals that the mean likelihood assessments in the partial-cue pattern condition (67.44) is greater than the mean in the no-cue pattern condition ($F_{1, 62} = 5.824$, one-sided $p = 0.008$). The mean of the mismatched participants' likelihood assessments in the full-cue pattern condition (54.67) is at least as great as the mean in the no- and partial-cue pattern conditions (54.57); however, this difference is not significant ($F_{1, 62} = 0.000$, one-sided

TABLE 3
H2: Likelihood Assessments
LS Mean (SE) [n] Cell

H2a contrast: A < (B + C)/2
H2b contrast: (D + E)/2 ≤ F

Panel A: Likelihood Assessments

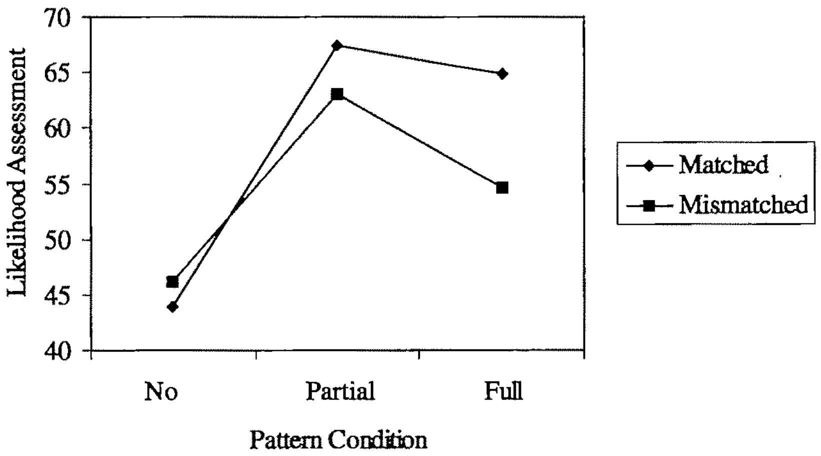
	Cue Pattern		
	No	**Partial**	**Full**
Matched	43.90	67.44	64.81
	(7.05)	(6.74)	(7.05)
	[21]	[23]	[21]
	A	B	C
Mismatched	46.14	63.00	54.67
	(6.23)	(5.95)	(6.23)
	[21]	[23]	[21]
	D	E	F

Panel B: Tests of H2

Planned Contrasts	**$F_{1, 62}$**	**p > F (one-tailed)**
Matched No < (Matched Partial + Full)/2	6.720	0.006
(Mismatched No + Partial)/2 ≤ Mismatched Full	0.000	0.495

See Table 2 for explanations of industry match and cue-pattern variables.
The dependent variable is the likelihood that a material misstatement was contained in the financial statements. Participants rate this likelihood on a 101-point scale anchored by 0 (No chance of material misstatement) and 100 (Certain of material misstatement).

FIGURE 3
Likelihood Assessment Results



See Figure 1 for explanations of industry match and cue-pattern variables.
The dependent variable is the likelihood that a material misstatement was contained in the financial statements. Participants rate this likelihood on a 101-point scale anchored by 0 (No chance of material misstatement) and 100 (Certain of material misstatement).

p = 0.495). Contrary to expectations, mismatched participants' likelihood assessments in the partial-cue pattern condition (63.00) are marginally higher than those in the no-cue pattern condition (46.14) ($F_{1,62}$ = 3.832, two-sided p = 0.055).

This evidence supports H2a for the matched participants; their likelihood assessments are higher in the full- and partial-cue pattern condition than in the no-cue pattern condition. However, H2b is not supported. Instead, it appears that mismatched participants may be aware that there is increased risk in the partial-cue pattern condition.

Hypothesis 3

In H3, I predict that matched specialists' problem representation development will mediate the relationship between likelihood assessments and the pattern-condition manipulation; this relationship is not predicted for mismatched specialists. I use mediation analysis to investigate this (Baron and Kenny 1986).[14]

Matched Specialists

I expect problem representation development to mediate the relationship between the independent variable, pattern completeness, and the dependent variable, matched likelihood assessment. The matched likelihood assessment is correlated with the problem representation measure ($r = 0.354$, Bonferroni adjusted $p = 0.038$).

---

[14] A mediation analysis requires a relationship to mediate. Generally, the following analysis is required: First, the dependent variable of interest (e.g., likelihood assessment) is shown to be related to the independent variable (e.g., pattern completeness). Second, the expected mediator variable (e.g., overall problem representation measure) is shown to be related to the independent variable. Third, the mediator variables are shown to be correlated with the dependent variable. Finally, the mediator and original independent variables are run against the dependent variable. If the mediator variable mediates the relationship between the dependent variable and the original independent variable, then the significance of the original independent variable will be reduced over the first ANOVA and the mediator variable will be significant in this analysis.

An ANOVA (Panel A of Table 4) shows that the matched likelihood assessment is related to the independent variable pattern completeness ($F_{2, 62} = 3.417$, $p = 0.039$). An ANOVA (Panel B of Table 4) shows that the problem representation measure is affected by pattern completeness ($F_{2, 62} = 4.226$, $p = 0.019$). Finally, an ANOVA (Panel C of Table 4) with matched likelihood assessment as the dependent variable and pattern completeness and the problem representation measure as independent variables shows that the relationship between the matched likelihood assessment and pattern completeness is mediated by problem representation development. In particular, problem representation development is significant ($F_{1, 59} = 5.634$, $p = 0.021$) and pattern completeness is no longer significant in this ANOVA ($F_{2, 59} = 1.666$, $p = 0.198$). This analysis demonstrates that pattern condition does not directly affect likelihood assessments for matched participants; instead, pattern condition affects matched likelihood assessments indirectly through its effect on problem representation development. This evidence supports H3.

Mismatched Specialists

The relationship examined, but not expected to mediate, is between the dependent variable mismatched likelihood assessment and the independent variable pattern condition and the potential mediator, the problem representation measure. The mismatched likelihood assessment is not correlated with the problem representation measure ($r = 0.066$, Bonferroni adjusted $p = 1.000$).

An ANOVA (not tabulated) shows that the mismatched likelihood assessment is not related to the independent variable pattern completeness ($F_{2, 62} = 1.917$, $p = 0.156$). Therefore, for the mismatched participants, the relationship examined between the dependent

TABLE 4
Matched Participants' Mediation Analysis

Source	Sum of Squares	df	F	p > F
Panel A: Effect of Cue Pattern on MatchLA				
Cue Pattern	7,130.28	2	3.417	0.039
Error	64,696.70	62		
Panel B: Effect of Cue Pattern on MatchPR				
Cue Pattern	50.10	2	4.226	0.019
Error	367.45	62		
Panel C: Effect of Cue Pattern and MatchPR on MatchLA				
Cue Pattern	3,264.97	2	1.666	0.198
MatchPR	5,520.90	1	5.634	0.021
Cue Pattern * MatchPR	1,517.32	2	0.774	0.466
Error	57,817.11	59		

Variable Definitions:
MatchLA = matched specialists' likelihood assessments; and
MatchPR = matched specialists' overall problem representation measures.
See Table 2 for explanation of the cue-pattern variable.

variable (likelihood assessments) and the independent variable (pattern completeness) is not significant and likelihood assessments are not correlated with the problem representation measure, so there is no potential for mediation, as expected.[15]

Hypothesis 4

In H4a, I predict that matched participants who receive full- and partial-cue patterns will allocate more time to procedures that discriminate whether a misstatement is present than will matched participants who receive no-cue patterns. In H4b, I predict that mismatched participants who receive full-cue patterns will allocate at least as much time to discriminatory procedures as will mismatched participants who receive no- or partial-cue patterns.

I estimate a repeated-measures ANOVA with matched and mismatched time allocated to discriminatory procedures as the dependent variable and industry match (within participant), case order, and cue-pattern condition (between-participant) as independent variables, and matched and mismatched total time allocated to all procedures as covariates. I present the adjusted cell means and results of planned contrasts in Panels A and B of Table 5 and the adjusted cell means are graphed in Figure 4. Matched participants' mean time allocated to discriminatory procedures in the partial- and full-cue pattern conditions (3.66) is greater than that in the no-cue pattern condition (1.69) and this difference is significant ($F_{1,57} = 3.389$, one-sided $p = 0.036$). Additionally, a *post hoc* test reveals that the mean time allocated to discriminatory procedures in the partial-cue pattern condition (3.04) is greater than the mean in the no-cue pattern condition, however, this difference is not significant ($F_{1,62} = 1.217$, one-sided $p = 0.137$). Mismatched participants in the full-cue pattern condition allocated at least as much time to discriminatory procedures (2.24) as those participants in the no- and partial-cue pattern conditions (1.35), ($F_{1,57} = 0.916$, one-sided $p = 0.171$). These results support H4a and H4b.

Hypothesis 5

In H5, I predict that mismatched participants will allocate more time to procedures that likely will *not* discriminate whether a misstatement is present than will matched participants. I estimate a repeated measures ANOVA with matched and mismatched time allocated to procedures that likely will not discriminate the presence of the target misstatement as the dependent variables and industry match (within participant), case order, and cue-pattern condition (between-participant) as independent variables, and matched and mismatched total time allocated to all procedures as covariates. Adjusted means and results of planned contrasts are presented in Panels C and D of Table 5. Mismatched participants allocated more time to procedures that did not discriminate whether a misstatement is present (8.88) than did matched participants (7.51) and this difference is significant ($F_{1,57} = 3.093$, one-sided $p = 0.042$). This evidence supports H5.

Additional Analysis

I perform additional analysis to ensure that participants performed better in their specialty industry due to knowledge differences rather than because they worked harder on the

---

[15] For completeness, an ANOVA shows that the problem representation measure is not affected by pattern completeness ($F_{2,62} = 1.077$, $p = 0.347$). Finally, an ANOVA with mismatched likelihood assessment as the dependent variable and pattern completeness and the problem representation measure as independent variables shows that the relationship between the mismatched likelihood assessment and pattern completeness is not mediated by the problem representation measure; pattern completeness is not significant in this ANOVA ($F_{2,59} = 1.395$, $p = 0.256$) and the problem representation measure also is not significant ($F_{1,59} = 0.166$, $p = 0.686$).

TABLE 5
Procedure Discrimination
Adjusted LS Mean (SE) [n] Cell

H4a contrast: A < (B + C)/2
H4b contrast: (D + E)/2 ≤ F

Panel A: Time Allocated to Discriminatory Procedures

	Cue Pattern		
	No	**Partial**	**Full**
Matched	1.69 (0.88) [21] A	3.04 (0.84) [23] B	4.28 (0.87) [21] C
Mismatched	1.06 (0.77) [21] D	1.63 (0.74) [23] E	2.24 (0.77) [21] F

Panel B: Tests of H4

Planned Contrasts	$F_{1,57}$	**p > F (one-tailed)**
Matched No < (Matched Partial + Full)/2	3.389	0.036
(Mismatched No + Partial)/2 ≤ Mismatched Full	0.916	0.171

Panel C: Time Allocated to Non-Discriminatory Procedures

	Cue Pattern		
	None	**Partial**	**Full**
Matched	8.77 (0.88) [21] A	7.42 (0.84) [23] B	6.19 (0.87) [21] C
Mismatched	9.43 (0.77) [21] D	8.86 (0.74) [23] E	8.25 (0.77) [21] F

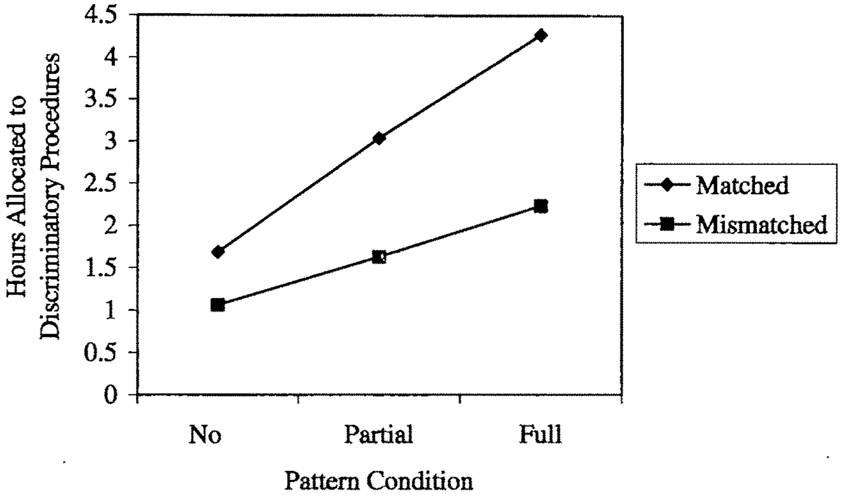
Panel D: Tests of H5

Planned Contrasts	$F_{1,57}$	**p > F (one-tailed)**
Mismatched > Matched	3.093	0.042

See Table 2 for explanations of industry match and cue-pattern variables.
The *dependent variable for H4a and H4b* is the total time, in hours, participants allocated to procedures that can discriminate whether the target misstatement was present. The *dependent variable for H5* is the total time, in hours, participants allocated to procedures that would not discriminate whether the target misstatement was present.

case that matched their specialization than on the case outside their specialization. Participants completed ratings of effort and difficulty immediately after the recall task for each case. They respond on 11-point Likert scales anchored by 0 (very easy) and 10 (very hard)

FIGURE 4
Discrimination of Procedures Results



See Figure 1 for explanations of industry match and cue-pattern variables.
The dependent variable is the total time participants allocated to procedures that can discriminate whether the target misstatement was present.

for the difficulty question and 0 (not at all hard) and 10 (very hard) for the effort question. Participants report that the cases mismatched on specialty are more difficult (6.01) than the cases matched on specialty (5.27); this difference is significant ($F_{1,60} = 10.501$, two-sided $p = 0.002$). They report that they work harder on the mismatched cases (5.78) than on the matched cases (4.17); this difference is significant ($F_{1,60} = 41.778$, two-sided $p = 0.000$).

Because participant self-reports of effort are noisy measures of effort (Bettman et al. 1990), I also compare participant work output (i.e., the number of items recalled for each case) as a measure of effort. The number of items recalled does not differ depending on whether the participants are matched on industry specialty ($F_{1,62} = 1.521$, two-sided $p = 0.222$). This analysis, in combination with the observed systematic differences in the problem representation development, provides evidence that knowledge, not effort, explains the results.

V. CONCLUDING REMARKS

Limitations

I am aware of two boundary conditions in relation to the theory I test in this paper. First, the theory applies to situations in which industry specialists are expected to have a comparative advantage over nonspecialists due to knowledge differences. This comparative advantage is expected to exist when patterns are composed of industry-specific cues or are focused on industry-specific transactions. Industry specialists are unlikely to have as large a comparative advantage in situations where there are opportunities for knowledge transfer by analogy or direct transfer. Knowledge transfer likely occurs when a situation in one industry is easily mapped into a target situation in another industry (Thibodeau 2003). Transfer is likely when auditing common "generic" accounts (e.g., property, plant, and equipment) as virtually all firms have such accounts and the audit procedures, accounting requirements, and sources of risk are similar across industries. Transfer also is likely in similar industries or where sources of risk, methods of accounting, or types of transactions are similar. In my study, to insure a powerful test of my theory, I chose industries that have little in common and I designed the cases around issues that are unique to these industries so that opportunities for knowledge transfer were minimized.

Second, the theory specifies how industry specialists are expected to react when a misstatement is present. It is likely that there are implications for how industry specialists react when a misstatement is absent as well; however, those predictions are not developed in this paper. This may be a fruitful area for future research.

Conclusions

Misstatements that are difficult to diagnose because they are complex or hidden from the auditor may be described by a pattern of cues. The division of labor on audits suggests the possibility that different auditors will collect the cues that form a pattern. Features of the audit ecology, such as the large amount of information to be processed, the types of tasks performed, and the size of audit teams make it likely that little sharing of unique information about pattern cues will occur. The brainstorming requirements of SAS No. 99 may improve unique information sharing for auditors in identified fraud risk areas. This is an open question that is left to future research. However, it remains important to understand how well individual auditors interpret incomplete patterns suggestive of misstatement.

In this paper, I examine whether industry-specialist auditors use their problem representations to help them interpret incomplete patterns that are descriptive of misstatement. Consistent with expectations, matched specialists develop more complete problem representations about a seeded misstatement when they receive partial- or full-cue patterns than when they receive no-cue patterns. These problem representations are related to differences in matched specialists' assessments of misstatement likelihood. Additionally, matched specialists specify audit procedures that effectively and efficiently discriminate the presence of the seeded misstatement. Mismatched specialists do not develop more complete problem representations with more pattern cues and their problem representations are not related to their assessments of misstatement likelihoods. Additionally, mismatched specialists do not suggest audit procedures that either would effectively or efficiently discriminate the presence of the seeded misstatement.

Overall, these results imply that matched specialists are able to interpret and fill in partial-cue patterns. They respond to partial-cue patterns by increasing the assessed risk of misstatement. Mismatched specialists do not suspect the seeded misstatement under these conditions. This suggests that identification of incomplete patterns of misstatement may be a comparative advantage for industry-specialist auditors.

My study contributes to the audit-knowledge literature by expanding our understanding of the conditions under which performance will improve due to industry specialization. Previous research has demonstrated that matched industry-specialist auditors know more non-error explanations for unusual fluctuations than mismatched specialists (Solomon et al. 1999). This gives rise to the concern that industry specialists may "explain away" anomalous facts as non-errors, rather than seriously consider them as indicators of misstatement. However, in a small sample study, Wright and Bedard (2000) report that auditors with high domain experience do concentrate on error hypotheses. I provide additional evidence that matched industry specialists develop more complete problem representations, use these representations to assess higher likelihoods of misstatement, and specify discriminatory procedures to determine whether the seeded misstatement exists.

My study contributes to the pattern-recognition literature by demonstrating that industry-specialist auditors are able to fill in the missing pieces of a partial-cue pattern. Previous research on pattern recognition studied auditors' ability to identify the significance of complete patterns indicative of misstatement leaving open the question of whether auditors identify the significance of incomplete patterns of misstatement. Interpreting incomplete patterns is arguably a more difficult task and one that is likely to assist in identification of misstatement concealed by a client.

APPENDIX A

State and Local Government Instrument[16]

City of Middletown Case

You have been assigned the revenue section of the general-purpose financial statement audit of the City of Middletown for the June 30, 2001 fiscal year. The City of Middletown has been a client of the firm for the past eight years. During that time there has been little turnover in the controller's office staff; the personnel are competent and it is easy to work with them. Adjustments, when necessary, have been made with little conflict.

The City's (unaudited) total revenue in the General Fund for the year ended June 30, 2001 is $25,211,000. Materiality has been set at $125,000 on this engagement.

You are currently focusing on the revenue in the Capital Projects Fund. Total (unaudited) revenue in the Capital Projects Fund is $904,000. The revenue comes from two sources: a *Federal Government Grant* for $250,000 and *Sewer Benefit Tax* revenue of $654,000.

The *Federal Government Grant* received during fiscal 2001 was for renovation of two park buildings and surrounding grounds. The grant period is January 1, 2001 to December 31, 2002, the period during which the improvements must be completed; any money unspent at the end of this period must be returned. The grant totals $3,420,000; the renovations to the buildings and grounds are expected to cost $4,275,000. The City has transferred the $855,000 of funds necessary to complete the project. They broke ground on the improvements in March 2001.

- The City has incurred $1,250,000 in costs on the project as of June 30; this is 29.24 percent of the project's current expected cost.[17]
- The City receives quarterly payments of $427,500 each on the 10th day of each calendar quarter; the City received payments this year on January 10 and April 10.
- The City uses a 60-day availability period for revenue recognition.

The *Sewer Benefit Tax* is collected from all city property owners to pay for sanitary sewer improvements. The sewer tax is billed annually and is based on the volume of water that is used and returned to the City's sanitary sewer system. Bills are sent on February 1 of each year, with payments due by March 15. Revenues from this tax are restricted as to use for payment for sanitary sewer improvements and maintenance. The City collected $654,000 and $621,000 in sanitary sewer tax during the fiscal years ended June 30, 2001 and 2000, respectively. The accounts receivable balances for uncollected Sewer Benefit Tax, net of estimated uncollectible balances, are $47,000 and $43,000 at June 30, 2001 and 2000, respectively.

---

[16] The issue in this case is how much of the Federal Government Grant revenue should be recognized during the June 2001 fiscal year. This reimbursement-driven grant is a Voluntary Nonexchange Transaction and "recipients should recognize revenues in the period when all applicable eligibility requirements have been met *and* the resources are available" (GASB No. 33, ¶30, GASB 1998). There are four eligibility requirements specified in the standard: required characteristics of recipients, time requirements, reimbursements, and contingencies (GASB No. 33, ¶20) and they have all been met by this grant. Three payments, or $1,282,500, of this grant are available as of June 30 because of the City's 60-day availability policy. The eligibility requirements limit the revenue recognition to the amount of reimbursable costs that have been incurred to date, here $1,250,000. Consequently, the client should recognize $1,250,000 of grant revenue in this case, representing a misstatement of $1,000,000.

[17] Note that this is the full-pattern condition government case. The partial-cue (no-cue) pattern condition case included the first two (none of the) bulleted points in the fourth paragraph.

APPENDIX B
Banking Instrument[18]

Springfield Bank Case

You have been assigned to the December 31, 2000 audit of Springfield Bank, a regional bank in Springfield, NJ. Springfield Bank has been a client of the firm for the past eight years. There has been little turnover in the accounting staff during this time period. The bank's management and accounting staff are competent, and it is easy to work with them. Additionally, on previous engagements management has readily agreed to make the audit adjustments the firm has determined to be necessary.

Current year (unaudited) interest revenue totals $82.6 million. Materiality for this engagement has been set at $400,000.

You are currently focusing on the fee revenue. Fee revenue for the current year totals $1,357,000 and comprises 14.5 percent of the bank's net income of $9,385,000. Fee revenue comes from two sources: *loan fees* of $910,000 and *credit card fees* of $447,000.

Loan fees are collected when a borrower applies for a mortgage; the fees are nonrefundable. During 2000, Springfield received about 24,200 mortgage loan applications. To originate a mortgage, Springfield incurs various costs. These include costs associated with the application, appraisal, credit report, underwriting, and closing. Springfield prices its loan fees to be competitive in its market.

- During 2000, Springfield closed approximately 23,000 mortgage loans.[19]
- The loan fee price of $375 covers the costs associated with originating a mortgage and provides a small surplus of about $20 per loan.
- When mortgage loans close, loan fee revenue equal to the cost associated with making the loans is deferred.

Credit card fee revenue totaled $447,000 during 2000. The fees cover servicing and maintaining credit card accounts and other services offered to the credit card holders. Credit card fees are billed annually to card holders on the anniversary of the dates on which they opened their accounts. Total fees billed to credit card customers during 2000 were $452,000. The fees are deferred and recognized on a straight-line basis over the period the fee entitles the cardholder to use the card. Deferred credit card fees, net of estimated uncollectibles, total $222,000 and $217,000 at December 31, 2000 and 1999, respectively.

---

[18] The issue in this case is how much of the loan fee revenue should be recognized during the year. These loan fees are direct loan origination fees and as such they should be deferred and recognized over the life of the loan as an adjustment to yield (interest income). Additionally, direct loan origination costs must be deferred and recognized as a reduction in loan yield. The loan fees and costs must be offset and the net amount should be deferred and amortized (SFAS No. 91, ¶5, FASB 1986). Consequently, loan fee revenue is overstated by $460,000 in this case.

[19] This is the full-pattern condition banking case. The partial-cue (no-cue) pattern condition case included the first two (none of the) bulleted points in the fourth paragraph.

REFERENCES

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Bedard, J. C., and S. F. Biggs. 1991a. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* 66 (3): 622–642.

———, and ———. 1991b. The effect of domain-specific experience on evaluation of management representations in analytical procedures. *Auditing: A Journal of Practice & Theory* 10 (Supplement): 77–90.

———, and A. M. Wright. 1994. The functionality of decision heuristics: Reliance on prior audit adjustments in evidential planning. *Behavioral Research in Accounting* 6 (Supplement): 62–89.

Bettman, J. R., E. J. Johnson, and J. W. Payne. 1990. A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes* 45 (1): 111–139.

Bierstaker, J. L., J. C. Bedard, and S. F. Biggs. 1999. The role of problem representation shifts on auditor decision processes in analytical review procedures. *Auditing: A Journal of Practice & Theory* 18 (1): 18–36.

Biggs, S. F., T. J. Mock, and P. R. Watkins. 1988. Auditors' use of analytical review and audit program design. *The Accounting Review* 63 (1): 148–161.

Brewer, W. F. 1987. Schemas versus mental models in human memory. In *Modelling Cognition*, edited by P. Morris, 189–197. Chichester, U.K.: Wiley.

Brown, C. E., and I. Solomon. 1990. Auditor configural information processing in control risk assessment. *Auditing: A Journal of Practice & Theory* 9 (3): 17–38.

———, and ———. 1991. Configural information processing in auditing: The role of domain-specific knowledge. *The Accounting Review* 66 (1): 100–119.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review* 65 (4): 933–945.

Carmichael, D. R. 1997. Report card on the accounting profession. *The CPA Journal* 67 (1): 18.

Christ, M. Y. 1993. Evidence on the nature of audit planning problem representations: An examination of auditor free recalls. *The Accounting Review* 68 (2): 304–322.

Dutke, S. 1986. Generic and generative knowledge: Memory schemata in the construction of mental models. In *Processes of the Molar Regulation of Behavior*, 35–54. Scottsdale, AZ: Pabst Science Publishing.

Erickson, M. M., B. W. Mayhew, and W. L. Felix, Jr. 1996. Understanding the client's business: Lessons from Lincoln Savings and Loan. Working paper, The University of Arizona.

Financial Accounting Standards Board (FASB). 1986. *Accounting for Nonrefundable Fees and Costs Associated with Originating or Acquiring Loans and Initial Direct Costs of Leases.* Statement of Financial Accounting Standards No. 91. Stamford, CT: FASB.

Fincher-Kiefer, R., T. A. Post, T. R. Greene, and J. F. Voss. 1988. On the role of prior knowledge and task demands in the processing of text. *Journal of Memory and Language* 27 (4): 416–428.

Governmental Accounting Standards Board (GASB). 1998. *Accounting and Financial Reporting for Nonexchange Transactions.* Governmental Accounting Standards Board Statement No. 33. Norwalk, CT: GASB.

Greeno, J. G., 1989. Situations, mental models, and generative knowledge. In *Complex Information Processing: The Impact of Herbert A. Simon*, edited by D. Klahr, and K. Kotovsky, 285–318. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnson-Laird, P. N. 1983. *Mental Models*. Cambridge, MA: Harvard University Press.

Kennedy, S. J., D. N. Kleinmuntz, and M. E. Peecher. 1997. Determinants of the justifiability of performance in ill-structured tasks. *Journal of Accounting Research* 35 (Supplement): 105–123.

Kerr, D. S., and D. D. Ward. 1994. The effects of audit task on evidence integration and belief revision. *Behavioral Research in Accounting* 6: 21–42.

Maletta, M. J., and T. Kida. 1993. The effect of risk factors on auditors' configural information processing. *The Accounting Review* 68 (3): 681–691.

Norman, D. A. 1983. Some observations on mental models. In *Mental Models*, edited by D. Gentner, and A. L. Stevens, 7–14. Hillsdale, NJ: Lawrence Erlbaum Associates.

Owhoso, V. E., W. F. Messier, and J. G. Lynch. 2002. Error detection by industry-specialized teams during sequential audit review. *Journal of Accounting Research* 40 (3): 883–900.

Payne, J. W., J. R. Bettman, and E. J. Johnson. 1993. *The Adaptive Decision Maker*. New York, NY: Cambridge University Press.

Rumelhart, D. E., and D. A. Norman. 1988. Representation in memory. In *Steven's Handbook of Experimental Psychology, Learning, and Cognition*, Vol. 2, edited by R. C. Atkinson, R. J. Herstein, G. Lindzey, and R. D. Luce, 511–587. New York, NY: Wiley.

Solomon, I., M. Shields, and O. R. Whittington. 1999. What do industry-specialist auditors know? *Journal of Accounting Research* 37 (Spring): 191–208.

Spilich, G. J., G. T. Vesonder, H. L. Chiesi, and J. F. Voss. 1979. Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior* 18 (3): 275–290.

Stasser, G., and W. Titus. 1985. Pooling of unshared information in group decision making: Biased information sampling during group discussion. *Journal of Personality and Social Psychology* 48 (6): 1476–1478.

———, and ———. 1987. Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology* 53 (1): 81–93.

———, L. Taylor, and C. Hanna. 1989. Information sampling in structured and unstructured discussions of three- and six-person groups. *Journal of Personality and Social Psychology* 57 (1): 67–78.

———, and D. Stewart. 1992. Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology* 63 (3): 426–434.

Tardieu, H., M. F. Ehrlich, and V. Gyselinck. 1992. Levels of representation and domain-specific knowledge in comprehension of scientific texts. *Language and Cognitive Processes* 7 (3–4): 335–351.

Taylor, M. 2000. Bounded rationality, uncertainty, and competence: The effects of industry specialization on auditors' inherent risk assessments and confidence judgments. *Contemporary Accounting Research* 17 (Winter): 693–712.

Thibodeau, J. 2003. The development and transferability of task knowledge. *Auditing: A Journal of Practice & Theory* 22 (3): 47–67.

Voss, J. F., G. T. Vesonder, and G. T. Spilich. 1980. Text generation and recall by high-knowledge and low-knowledge individuals. *Journal of Verbal Learning and Verbal Behavior* 19 (6): 651–667.

Wright, A. M., and J. C. Bedard. 2000. Decision processes in audit evidential planning: A multistage investigation. *Auditing: A Journal of Practice & Theory* 19 (1): 123–143.

Wright, S., and A. M. Wright. 1997. The effect of industry experience on hypothesis generation and audit planning decisions. *Behavioral Research in Accounting* 9: 273–294.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 337–375

Risk-Relevance of Fair-Value Income Measures for Commercial Banks

Leslie D. Hodder
Indiana University

Patrick E. Hopkins
Indiana University

James M. Wahlen
Indiana University

ABSTRACT: We investigate the risk relevance of the standard deviation of three performance measures: net income, comprehensive income, and a constructed measure of full-fair-value income for a sample of 202 U.S. commercial banks from 1996 to 2004. We find that, for the average sample bank, the volatility of full-fair-value income is more than three times that of comprehensive income and more than *five times* that of net income. We find that the incremental volatility in full-fair-value income (beyond the volatility of net income and comprehensive income) is positively related to market-model beta, the standard deviation in stock returns, and long-term interest-rate beta. Further, we predict and find that the incremental volatility in full-fair-value income (1) negatively moderates the relation between abnormal earnings and banks' share prices and (2) positively affects the expected return implicit in bank share prices. Our findings suggest full-fair-value income volatility reflects elements of risk that are not captured by volatility in net income or comprehensive income, and relates more closely to capital-market pricing of that risk than either net-income volatility or comprehensive-income volatility.

We gratefully acknowledge helpful comments and suggestions from T. Shevlin, three anonymous reviewers, M. Barth, W. Beaver, W. Blacconiere, J. Hand, E. Hirst, W. Landsman, R. Lipe, L. Maines, R. Martin, C. Nichols, S. Ryan, G. Salamon, L. Vincent, B. Walther, W. Wallace, participants at The Frank Batten Young Scholars Forum in Accounting at the College of William & Mary, the Stanford Summer Camp, the UNC/Duke Fall Camp, the Minnesota Empirical Accounting Research Conference, and workshop participants at Carnegie Mellon University, Indiana University, University of Nebraska, Northwestern University, University of Oklahoma, and the Federal Reserve Bank of St. Louis. We thank Craig Nichols for his research assistance. We also thank I/B/E/S for analyst forecast data and SNL Financial for fair-value footnote disclosures. We gratefully acknowledge the financial support of the Kelley School of Business and the Graduate School of Business at Stanford University. Professor Hopkins and Professor Wahlen thank BKD LLP and the Ford Motor Company, respectively, for their financial support.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted June 2003
Accepted August 2005

I. INTRODUCTION

Under current United States generally accepted accounting principles (GAAP) and International Financial Reporting Standards (IFRS), net income and comprehensive income include unrealized fair-value gains and losses on some financial instruments but exclude them on the majority of financial instruments. In recent years, U.S. and international accounting rule makers have discussed whether to require income recognition of fair-value gains and losses on a more complete set of financial and derivative-financial instruments. Net income, comprehensive income, and "full fair-value income" (FFV income) can portray firm performance and risk differently for firms with significant exposure to changes in fair values of financial instruments.

Commercial banks' balance sheets consist almost entirely of financial instruments. The current income-measurement regime produces reported-income numbers that exclude unrealized fair-value gains and losses on financial instruments that are integral to banks' core operations and risk-management strategies, including held-to-maturity securities, loans, deposits, other financial liabilities, and certain types of derivatives. Given the importance of these financial instruments and derivatives for assessing bank performance and risk, we construct a measure of FFV income that includes unrealized fair-value gains and losses on all of these financial instruments and derivatives for a sample of U.S. commercial banks. We then address three research questions on the risk-relevance of the time-series volatilities of banks' net income, comprehensive income, and our measure of FFV income. The objective of the paper is to evaluate how well these three measures of income volatility represent commercial bank risk.

Our first research question asks: How do the volatilities of these three income measures compare? These comparisons are interesting because, during the last decade, bank managers and their representatives have made conflicting assertions regarding the usefulness of income measures that include changes in fair values of financial instruments and derivatives. For example, in the mid-1990s banking organizations criticized the Financial Accounting Standards Board's (FASB) proposal for comprehensive-income measurement and reporting because it would fail to fully reflect banks' risk-management activities, trigger "excessive" volatility in comprehensive income, and lead investors to overstate their risk assessments (Hirst et al. 2002). In striking contrast to their prior statements, these same organizations recently defended the recognition of fair-value changes on a subset of financial instruments in the currently reported measure of comprehensive income as being the "optimal means of reporting financial performance" for banks. They asserted that FFV accounting will provide a "false and misleading picture" of excessive volatility in banks' profitability (Dean 2000, 3) and will misrepresent banks' risk because it is irrelevant and unreliable (Joint Working Group of Banking Associations on Financial Instruments 1999).

If banks effectively hedge at an entity-wide level the risk factors that determine fair-value gains and losses on financial instruments, and if comprehensive income recognizes fair-value gains and losses on only a subset of financial instruments, then FFV income may be less volatile over time than comprehensive income. In contrast, if comprehensive income only partially captures banks' exposures to fair-value gains and losses, then FFV income may be more volatile over time than comprehensive income. In either case, we predict that FFV income is a more complete representation of the outcomes of banks' risk-management activities; therefore, volatility in the time-series of FFV income should be more indicative of banks' underlying risks.

Our second research question addresses whether the three different measures of income volatility are associated with market-based risk measures. To address this question, we examine the associations between the three income-volatility measures and four stock-market-based risk factors (i.e., market-model beta, short-term and long-term interest-rate betas, and stock-return volatility), as well as two disclosed measures of banks' exposure to market-based risks (i.e., total exposure to derivatives and the gap between fixed-rate interest-bearing assets and fixed-rate interest-bearing liabilities that reprice within one year). To the extent these market-based risk measures represent systematic- and total-risk factors, our analysis provides evidence on the risk-related information content of each income-volatility measure. If net income and comprehensive income do not fully reflect the fair-value gains and losses that are triggered by banks' overall risk exposure, then volatility based on these two income measures will misrepresent banks' risks. We predict that, because FFV income includes fair-value changes for nearly all financial instruments, volatility in FFV income should be a more complete proxy for systematic and total risk, and therefore should be more closely associated with banks' market-based risk measures than either volatility in net income or comprehensive income.

Our third research question investigates the extent to which the incremental components of the three income-volatility metrics moderate the capitalization of earnings in banks' share prices and explain capital-market pricing of bank risk. If the income-volatility metrics capture risk factors that the capital markets price, then higher volatility should be associated with greater risk and therefore higher expected returns, which should result in lower capitalization of earnings in share price, all else equal. We address this question in two stages. First, we examine the extent to which cross-sectional variation in abnormal-earnings-valuation multiples is conditional on each of the income-volatility metrics within a residual-income-valuation model. We predict that abnormal-earnings-valuation multiples implicit in bank share prices are more negatively related to volatility in our unreported-but-more-complete FFV-income measure, and less negatively related to volatility in reported-but-incomplete net income and comprehensive income. Second, we examine the extent to which expected returns implicit in bank share prices are conditional on each of the income-volatility metrics. We predict that bank share prices reflect required rates of return on bank equity that are increasing in the incremental volatility in FFV income, beyond the volatility in net income and comprehensive income.

To address our research questions we construct a panel of 202 U.S. commercial banks with complete fair-value information and stock-price data for the years 1996–2004. For each bank, we obtain three measures of income—reported net income, reported comprehensive income, and our measure of FFV income, which recognizes fair-value gains and losses on virtually all financial instruments in income in the periods when they arise. We estimate volatility by calculating the standard deviation of the time-series of each of the three income measures (as a percent of average total assets) for each bank.

With respect to our first question, we find that FFV-income volatility is significantly greater than net-income volatility for 90 percent of our sample banks, and significantly greater than comprehensive-income volatility for 77 percent of our sample banks. For the majority of our sample banks, the fair-value adjustments banks recognize in comprehensive income do not covary significantly negatively with (i.e., do not appear to hedge) the incremental fair-value adjustments that banks would recognize in FFV income. Further, the volatility of the incremental FFV-income adjustments greatly exceeds the volatility of fair-value adjustments recognized in comprehensive income.

Regarding our second research question, we find that the volatilities of all three income measures exhibit varying degrees of positive correlation with market-risk factors and disclosed measures of market-risks. In particular, net income volatility exhibits the most consistent and robust correlations across the risk factors we examine, and correlates most strongly with the standard deviation in stock returns. We find that FFV-income volatility correlates positively with the standard deviation in stock returns and banks' exposure to derivatives. We also find that the incremental volatility in FFV income (beyond volatility in net income and comprehensive income) is positively associated with the standard deviation in stock returns, market-model beta, and long-term interest-rate beta, even after controlling for banks' disclosed measures of total exposure to derivatives and the gap between fixed-rate interest-bearing assets and fixed-rate interest-bearing liabilities that reprice within one year.

Finally, to address our third research question, we predict and find that the incremental volatility in FFV income negatively moderates the capitalization of abnormal earnings in bank share prices. We also predict and find that expected returns on bank equity are increasing in the incremental volatility in FFV income. We observe these results in bank share prices and expected rates of return after controlling for the effects of the incremental volatilities in net income and comprehensive income and after controlling for the information in banks' disclosures about exposure to market risks related to the asset-liability repricing gap and derivatives.

Taken together, our results suggest that the majority of our sample banks are not fully hedged against year-to-year changes in fair values of their reported financial instruments, and that volatility in net income and comprehensive income provide an incomplete picture of their fair-value-risk exposure. The results show that, for our sample of banks during our study period, the volatility of incremental FFV income captures elements of bank risk that the capital markets price, but that the volatilities of net income and incremental comprehensive income omit. We believe these findings are relevant for U.S. and international accounting standard setters as they consider whether to develop standards that would recognize in income fair-value changes for all financial instruments.[1] These findings are also useful to capital-market participants and researchers interested in explaining the relations among banks' accounting numbers, share prices, and risk. The findings in this paper are also useful for bank regulators as they evaluate the role of fair-value measurement in monitoring bank capital adequacy. Academic research has established the risk-relevance of net-income volatility (e.g., Beaver et al. 1970), the value-relevance of fair values for a subset of financial instruments (e.g., Barth 1994), and the risk-relevance of market risk-related disclosures (e.g., Schrand 1997). We extend these three lines of research by demonstrating the risk-relevance of FFV-income volatility.

We organize the paper in the following sections. In Section II, we motivate our research questions with the institutional features of fair-value accounting for banks and the prior research literature. In Section III, we describe our sample selection procedures and our variables. In Section IV, we describe three sets of tests and present the results. In Section V, we provide concluding remarks.

II. INSTITUTIONAL BACKGROUND AND LITERATURE REVIEW

Income measurement is a central function of accounting and is controversial. Current and proposed U.S. and international accounting standards include (at least) two different measures of income—net income and comprehensive income—each including unrealized fair-value gains and losses on different subsets of financial instruments (e.g., Statement of Financial Accounting Standards (SFAS) No. 130, FASB 1997). The Joint Working Group of Standards Setters (JWGSS) was formed in 1997 for the sole purpose of developing a coherent framework for reporting financial instruments at fair value. Although it is not a standard-setting body, the JWGSS developed a proposal to require full fair-value accounting for most financial instruments (JWGSS 2000). The FASB and the International Accounting Standards Board (IASB) are considering requiring income-statement recognition of fair-value gains and losses on *all* financial instruments, which would result in a FFV-income measure and the most complete fair-value accounting standards to date.[2]

---

[1] We cannot generalize these finding to settings in which full fair-value income measurement is mandatory and bank managers face explicit contracting and regulatory constraints that depend on full fair-value income numbers. We do not purport that these results demonstrate the desirability of full fair-value accounting. Our evidence on the risk-relevance of full fair-value income volatility is only one of many facets of full fair-value accounting that standard setters might weigh in their deliberations. For a more complete debate and critique of the strengths and limitations of drawing inferences for accounting standard setters from capital-markets-based accounting research, see Holthausen and Watts (2001) and Barth et al. (2001).

The banking industry is in the center of the controversy over the relative usefulness of fair-value-based income measures because commercial banks' balance sheets consist almost entirely of financial instruments. For banks, these three different income metrics could provide strikingly different pictures of profitability and income volatility over time. For example, net income includes many (but not all) effects of core-banking activities, including interest revenue and expense, credit losses (i.e., loan-loss provisions), realized gains and losses on investments, operating expenses, fees, and taxes. Comprehensive income is a more complete measure of performance than net income because it also includes fair-value gains and losses on available-for-sale investment securities.[3] However, comprehensive income is an incomplete measure of bank performance because it does not include fair-value gains and losses on financial instruments such as held-to-maturity securities, loans, financial liabilities, and non-term deposits.[4] These additional fair-value gains and losses are important elements of bank risk and profitability, and would be included in measures of FFV income.

Our paper advances three lines of academic research investigating the relations between reported income measures, fair values, risk, and share values. First and most generally, for a sample of non-bank firms, Beaver et al. (1970) establish the risk-relevance of earnings volatility and document that earnings volatility and earnings beta explain market model beta. More recent studies establish links between accounting-based risk measures and the pricing of risk in expected returns and share prices.[5]

Second, studies based on SFAS Nos. 107 and 115 (FASB 1991, 1993) document the value-relevance of fair values for certain balance-sheet components, but obtain mixed results on the association between fair-value gains and losses and stock returns (e.g., Barth 1994).[6] Third, recent studies examine the risk-relevance of various disclosures, such as interest-rate gap, notional amounts of derivatives, and the Securities and Exchange Commission's (SEC)

---

[2] In this paper, our references to FFV accounting are consistent with the financial-instruments focus of the JWGSS proposal.

[3] Under current accounting rules, reported comprehensive income also reflects the effects of other items, specifically unrealized gains and losses on foreign-currency translation adjustments, minimum-pension liability adjustments, and unrealized gains and losses from derivatives qualifying as cash-flow hedges. For the average bank, these items are typically less material than unrealized gains and losses on available-for-sale securities.

[4] Prior to application of SFAS No. 133 (i.e., fiscal years ending before 2001, FASB 1998), net income and comprehensive income also exclude fair-value gains and losses on derivative positions a bank may hold to manage interest-rate risk.

[5] For more recent examples, see Elgers (1980), Fama and French (1992), Feltham and Ohlson (1999), Claus and Thomas (2001), Gebhardt et al. (2001), Botosan and Plumlee (2002, 2005), Easton et al. (2002), and Baginski and Wahlen (2003). For a comprehensive review of the literature linking accounting numbers and risk, see Ryan (1997).

[6] For additional examples, see Petroni and Wahlen (1995), Ahmed and Takeda (1995), Barth et al. (1996), Eccher et al. (1996), Nelson (1996), Dhaliwal et al. (1998), and Cahan et al. (2000).

Financial Reporting Release No. 48 market-risk disclosures, which relate to fair-value-income volatility (e.g., Schrand 1997).[7] Our study links and extends these three prior lines of research in that this is the first study to use a comprehensive set of fair-value gains and losses to construct full fair-value income, and to assess the risk-relevance of different measures of income volatility.

Our study most closely relates to Barth et al. (1995). Using a 1971–1990 sample period, which precedes the effective dates for SFAS No. 107 and SFAS No. 115, Barth et al. (1995) examine the incremental volatility in income induced by including unrealized fair-value gains and losses on all investment securities disclosed in bank regulatory reports. Barth et al. (1995) find that volatility in their constructed, partial-fair-value income metric is significantly greater than volatility in reported net income, but that the incremental volatility in partial-fair-value income is not priced as risk. They point out that their results may be attributable to incomplete measurement of fair values in income, because banks did not disclose fair-value gains and losses on financial instruments other than investment securities.

Our study also closely relates to prior experimental research in Hirst et al. (2004), which investigates the effects of alternative fair-value income measures for banks, and finds that risk assessments and valuation judgments of bank analysts are more likely to distinguish between high and low risk banks when banks measure and recognize *all* fair-value gains and losses in a FFV income statement. Because Hirst et al. (2004) limited their investigation to analysts' judgments, they do not make inferences about the relation between fair value information and market-determined security prices.

III. SAMPLE AND DESCRIPTIVE STATISTICS

In this section we describe our sample selection procedures, variables, and tests. We also discuss descriptive statistics for our sample data.

Sample Selection

We draw our sample from the set of U.S. bank holding companies filing Y9-C reports with the Federal Reserve and financial statements with the SEC from 1995–2004. We collect accounting data, including fair-value footnote disclosure data, from Y9-C reports and the SNL DataSource database. When fair-value data items are missing from these sources, we hand-collect them directly from banks' Forms 10-K and Y9-C reports. We collect stock prices and returns from CRSP. Our final sample comprises the set of 202 banks for which we could collect complete data for the ten years 1995–2004.[8] We use 1995 data for differencing, leaving nine years of data to calculate firm-specific standard deviations of income and other variables of interest. In total, we have a sample of 1,818 bank-years for our tests. Because of these data requirements, our sample inevitably excludes banks that were acquired, failed, started up, or went public during our sample period, and banks that do not file Form 10-K (typically smaller, privately held banks).

Table 1 provides descriptive statistics that indicate that our sample banks represent a broad cross-section, with market value of equity over the nine-year period ranging from $2 million to $190 billion, and average total assets ranging from $41 million to $1,045 billion. For comparison, we examined descriptive statistics (untabulated) for the sample of banks

---

[7] In a related paper, Hodder (2005) documents the value-relevance of banks' market risk disclosures based on cash flow and fair-value volatilities.

[8] For the subset of our tests that examine expected rates of return implicit in bank share prices, we also collect consensus analyst earnings forecast data from the Institutional Brokerage Estimates System (I/B/E/S). Because not all sample banks are followed by analysts that report their forecasts to I/B/E/S, this data requirement further restricts our sample for these tests to the subset of (larger) sample banks with available forecast data.

TABLE 1
Descriptive Statistics for Sample Banks

Sample of 202 banks with complete data over 1996–2004, pooled across years.
(n = 1,818 bank/years; amounts in $ millions except for ratio variables.)

	Mean	Std. Dev.	Min	1st Quartile	Median	3rd Quartile	Max
Average total assets	15,745.66	71,958.89	40.73	552.69	1,369.27	5,636.09	1,044,660.00
Total loans	7,066.22	32,562.48	0.00	79.66	581.35	2,428.05	521,837.00
Total loans/Average total assets	0.51	0.32	0.00	0.19	0.64	0.73	1.31
Total deposits	9,681.10	39,180.92	0.00	450.07	1,078.83	4,075.98	618,570.00
Total deposits/Average total assets	0.79	0.14	0.00	0.72	0.81	0.88	1.54
Book value of equity	1,294.14	5,657.96	1.32	51.28	123.90	502.64	105,314.00
Market value of equity	2,882.27	11,222.45	2.41	74.53	228.15	967.66	189,800.93

that met our accounting and stock-price data requirements for at least one year during 1996–2004. In general, for the variables of primary interest, our sample appears representative of the larger population of publicly held banking firms, diminishing concerns about survivorship and size bias.

Income Measures

We construct our measure of FFV income to be as complete a measure of FFV income as possible, encompassing fair-value accounting for all financial instruments. We compute our measure of FFV income by adjusting comprehensive income for unrealized fair-value gains and losses disclosed under SFAS Nos. 107 and 133 for held-to-maturity investment securities, loans, term deposits, long-term financial liabilities, and derivatives. Using Form Y9-C regulatory report data, we also adjust comprehensive income for unrealized fair-value gains and losses on non-term deposits. Fair values of non-term deposits are particularly difficult to estimate. We estimate their fair values using a static discounted cash flow approach used by bank regulators, together with firm-specific deposit data obtained from bank holding company Y9-C reports. Finally, to be consistent with the after-tax measures of comprehensive income and net income, we measure all of our fair-value adjustments after tax effects using the statutory Federal income tax rate. We describe our variable computations in detail (particularly the fair-value computations for non-term deposits) in the Appendix.

Table 2 contains descriptive statistics for the three primary income measures and selected components of income. We present the variables in the order in which the income components and totals might appear in a typical bank income statement if the bank were to report FFV income. For example, the panel begins with net interest income, the least inclusive income measure, and concludes with FFV income, the most inclusive. We express all variables as a percentage of average total assets each year to allow comparison across size within our cross-section of banks.[9] Panel A contains descriptive statistics for the full sample of 202 banks over 1996–2004 (i.e., n = 1,818). Panel B contains annual means for the full sample for each year 1996–2004.

The statistics in Panel A of Table 2 indicate that the pooled nine-year means (medians) of these three income measures have statistically indistinguishable magnitudes as a percent of average total assets: 1.133 (1.136) percent for net income versus 1.149 (1.172) percent for comprehensive income versus 1.026 (1.069) percent for FFV income (test results not tabulated). The descriptive statistics in Table 2, Panel B reveal significant differences across income measures in certain years. For example, for the mean bank in 1998 and 2000, years in which interest rates fell, comprehensive income as a percent of average total assets was much higher than full fair-value income as a percent of average total assets, because FFV income included the fair-value losses on financial liabilities, derivatives, and non-term deposits, which offset fair-value gains on financial assets.[10] In 1999, by contrast, a year in which interest rates jumped, the average bank experienced large unrealized fair-value losses

---

[9] We do not adjust average total assets for FFV gains and losses. This choice increases the FFV-income volatility metrics but only very slightly because net unrealized fair-value gains and losses tend to be a very small percentage of total assets for the typical bank, even though they can be large relative to net income.

[10] To gauge the economic significance of the reported differences in income measures when scaled by average total assets, consider that the average total assets for the mean bank-year in our sample is $15,745.66 million. Therefore, a 0.10 percent difference in income scaled by average total assets equates to a $15.746 million difference in income for the mean bank-year. Also, because banks are typically highly leveraged, a 0.10 percent difference in income scaled by average total assets equates to a 1.22 percent difference in return on equity for the mean bank-year.

TABLE 2
Descriptive Statistics for Income Measures and Selected Components of Income

Panel A: Descriptive statistics for the sample of 202 banks with complete data over 1996–2004, pooled across banks and years (n = 1818). We highlight in bold the aggregated income subtotals; income components are inset and in normal font. We scale each variable as a percentage of average total assets each year.

	Mean	S.D.	Min	1st Quartile	Median	3rd Quartile	Max
Net interest income	3.744	0.894	0.357	3.233	3.704	4.222	7.615
Provision for loan losses	0.266	0.286	−0.663	0.124	0.214	0.333	4.109
Realized gain (loss) on investment securities	0.035	0.088	−0.714	0.000	0.010	0.051	0.818
Net income	**1.133**	**0.499**	**−2.791**	**0.888**	**1.136**	**1.408**	**3.846**
Unrealized gain (loss) on available-for-sale securities	0.003	0.293	−2.171	−0.099	0.007	0.128	1.997
Comprehensive income	**1.149**	**0.584**	**−3.905**	**0.843**	**1.172**	**1.473**	**3.681**
Unrealized gain (loss) on financial assets	0.036	1.220	−23.396	−0.359	0.009	0.393	23.870
Unrealized gain (loss) on financial liabilities	0.007	1.175	−22.351	−0.196	0.000	0.190	30.912
Unrealized gain (loss) on derivative contracts	−0.090	2.061	−70.892	0.000	0.000	0.000	14.931
Unrealized gain (loss) non-term deposits	−0.076	0.362	−1.690	−0.268	−0.122	0.106	2.154
Subtotal—Fair-value accounting adjustments	−0.123	2.579	−70.645	−0.553	−0.105	0.391	28.690
Full fair-value income	**1.026**	**1.623**	**−69.190**	**0.438**	**1.069**	**1.697**	**29.398**

Panel B: Annual means for the sample of 202 banks with complete data over 1996–2004, by year. We highlight in bold the aggregated income subtotals; income components are inset and in normal font. We scale each variable as a percentage of average total assets each year.

	1996	1997	1998	1999	2000	2001	2002	2003	2004
Net interest income	3.920	3.847	3.808	3.821	3.784	3.728	3.771	3.537	3.483
Provision for loan losses	0.247	0.253	0.262	0.254	0.263	0.329	0.338	0.278	0.170
Realized gain (loss) on investment securities	0.025	0.030	0.052	0.009	−0.002	0.046	0.054	0.066	0.031
Net income	**1.103**	**1.116**	**1.120**	**1.116**	**1.114**	**1.096**	**1.220**	**1.168**	**1.145**
Unrealized gain (loss) on available-for-sale securities	−0.048	0.101	0.030	−0.463	0.314	0.125	0.204	−0.139	−0.100
Comprehensive income	**1.124**	**1.287**	**1.150**	**0.656**	**1.421**	**1.220**	**1.418**	**1.019**	**1.045**
Unrealized gain (loss) on financial assets	−0.088	0.227	0.274	−0.769	0.267	0.736	0.240	−0.094	−0.473
Unrealized gain (loss) on financial liabilities	0.058	−0.045	−0.047	0.348	−0.194	−0.203	−0.362	0.107	0.404
Unrealized gain (loss) on derivative contracts	0.052	−0.088	−0.341	−0.107	0.080	−0.542	−0.006	0.122	0.022
Unrealized gain (loss) non-term deposits	0.102	−0.114	−0.380	0.599	−0.529	0.192	−0.151	−0.168	−0.232
Subtotal—Fair-value accounting adjustments	0.124	−0.021	−0.494	0.072	−0.377	0.184	−0.278	−0.033	−0.279
Full fair-value income	**1.248**	**1.266**	**0.691**	**0.728**	**1.045**	**1.649**	**1.151**	**0.869**	**0.871**

on available-for-sale investment securities, so comprehensive income (0.656 percent of average total assets) was much lower than net income (1.116 percent of average total assets). That year, the mean sample bank also experienced additional substantial fair value losses on financial assets, but these were partially offset by fair-value gains on financial liabilities and non-term deposits, so FFV income (0.728 percent of average total assets) presents a more favorable profit picture than comprehensive income.

Figure 1 graphically depicts mean net income, comprehensive income, and FFV income as a percentage of average total assets for the sample of banks each year and pooled across time, together with percentage changes in interest rates during the sample period. The time-series of mean net income is the smoothest of the three income series, while comprehensive income and FFV income exhibit much wider year-to-year changes. However, as noted above, the pooled nine-year means of these three income measures are similar. These data show that, in any given year, exposure to interest-rate changes and other risks triggered dramatic differences in the three income measures; but, on average, fair-value adjustments in both comprehensive income and FFV income tended to zero out over this period. Comparing the plots to percentage changes in interest rates, net income appears relatively *invariant* to rate changes. In contrast, comprehensive income *covaries negatively* with interest-rate changes and FFV income *covaries positively* with interest-rate changes.

IV. EMPIRICAL TESTS AND RESULTS

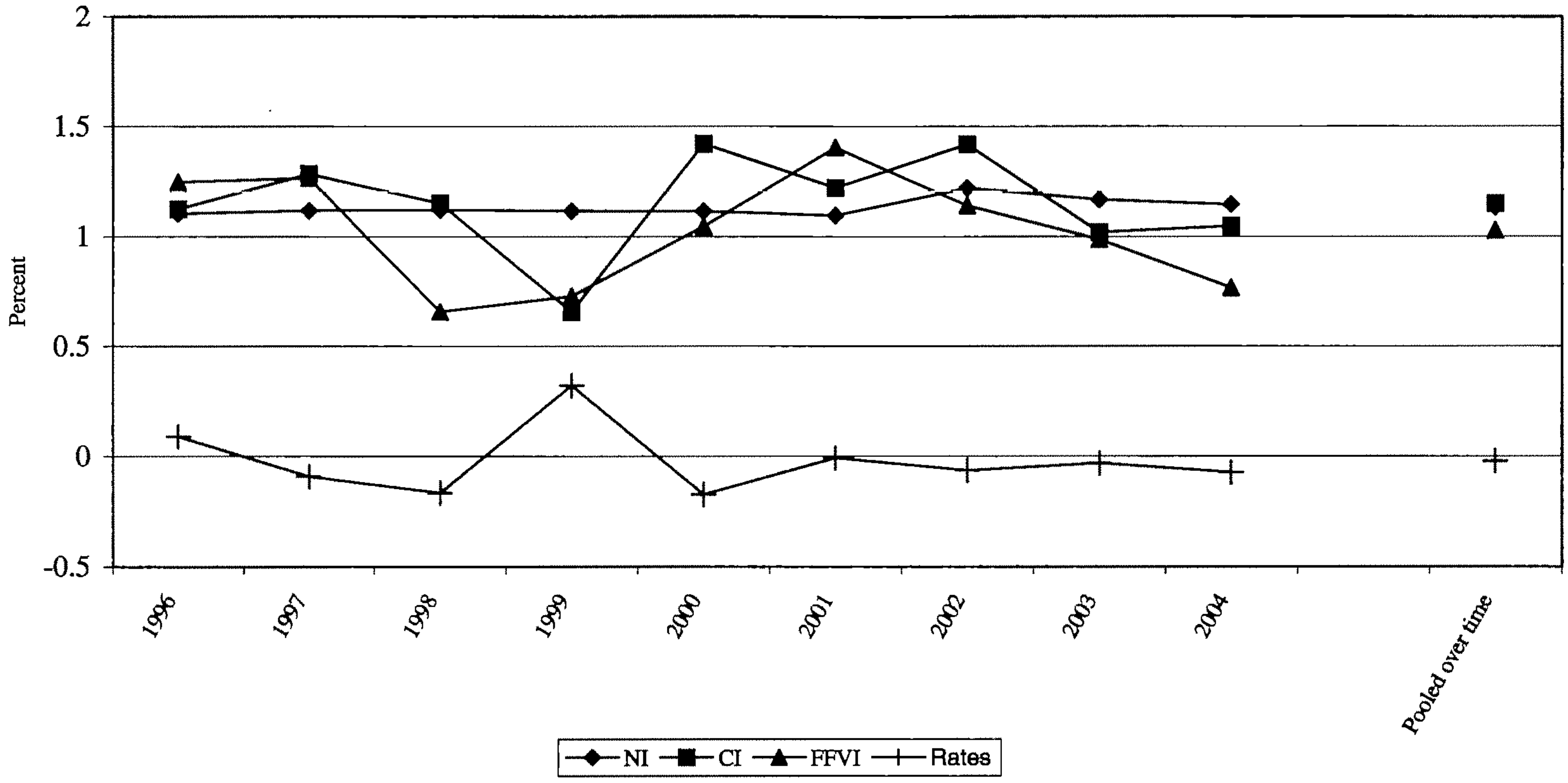
In this section, we describe our empirical tests and results. The section contains three parts, corresponding with our three research questions.

Income Volatility

The pattern of differences in volatility across income measures based on pooled data in Figure 1 and Table 2 is also evident in the firm-specific volatility analysis in Table 3. Panel A of Table 3 provides descriptive statistics for the firm-specific standard deviations of the three income measures (stated as a percentage of average total assets) across the 202 sample banks over the years 1996–2004. The average sample bank experienced a standard deviation in net income of 0.260 percent and a standard deviation in comprehensive income of 0.394 percent. The standard deviation in FFV income for the average sample bank was 1.442 percent, more than five times larger than the standard deviation in net income and more than three times larger than the standard deviation in comprehensive income. The left side of Figure 2 graphically depicts the means of the firm-specific standard deviations of our three income measures. The remainder of Figure 2 includes annual cross-sectional standard deviations (from Table 2, Panel B) and the pooled cross-sectional standard deviations (from Table 2, Panel A), which show that the standard deviation in FFV income is greater than that in comprehensive income and net income in each year, 1996 through 2004.

Panel B of Table 3 reports the results of firm-specific tests of the relative volatility of our three income measures. The first row compares the standard deviations of net income and comprehensive income. Of the 202 sample banks, 186 (92.1 percent) experienced greater volatility in comprehensive income than net income. For 69 sample banks (34.2 percent), firm-specific F-statistics reject ($p < 0.05$) the null hypothesis of equal variance (median F-statistic = 2.2). The second row of Panel B compares the standard deviations of net income and FFV income. The results indicate that FFV income is more volatile than net income for 200 sample banks (99.0 percent). For 182 banks (90.1 percent), firm-specific F-statistics reject ($p < 0.05$) the null hypothesis of equal variance (median F-statistic = 21.7). In the last row of Panel B, we compare standard deviations of FFV income and comprehensive income. FFV income is more volatile than comprehensive income for

FIGURE 1
Time-Series of Interest Rate Changes and Bank Income Measures



NI = cross-sectional mean net income as a percentage of average total assets each year;
CI = cross-sectional mean comprehensive income as a percentage of average total assets each year;
FFVI = cross-sectional mean full fair-value income as a percentage of average total assets each year;
Rates = the percentage change in spot treasury rates averaged over yield curve maturity intervals; and
n = 202 sample banks.

TABLE 3
Descriptive Statistics and Comparative Analyses for Measures of Income Volatility

Panel A: Descriptive Statistics for Firm-Specific Measures of Income Volatility over the Period 1996–2004

Standard Deviation of	Mean	Std. Dev.	Min	1st Quartile	Median	3rd Quartile	Max
Net income (NI)	0.260	0.199	0.045	0.141	0.201	0.308	1.471
Comprehensive income (CI)	0.394	0.221	0.119	0.258	0.346	0.467	2.002
Full fair-value income (FFVI)	1.442	2.220	0.257	0.677	0.961	1.471	25.302

Panel B: Firm-Specific Tests to Compare Measures of Income Volatility over the Period 1996–2004

			Folded F-Statistics (i.e., Levene's Test) for Null Hypothesis of Equality of Standard Deviations		
Comparisons	Count	Percent	Number that Reject Equality > <	Percent that Reject Equality > <	Median F-Statistic
$\sigma_{CI} > \sigma_{NI}$	186	92.1	69	34.2	2.2
$\sigma_{FFVI} > \sigma_{NI}$	200	99.0	182	90.1	21.7
$\sigma_{FFVI} > \sigma_{CI}$	194	96.0	155	76.7	7.6

Panel C: Descriptive Statistics for Firm-Specific Correlation Coefficients of Major Categories of Income Components over the Period 1996–2004

	Negative Count	Positive Count					
Pearson Correlations of	Signif./Total	Signif./Total	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
NI with CI-Adjustments	6/105	2/97	−0.024	0.324	−0.249	−0.013	0.169
CI with FFV-Adjustments	6/98	6/104	−0.003	0.389	−0.342	0.026	0.279
CI-Adjustments and FFV-Adjustments	15/100	7/102	−0.009	0.413	−0.314	0.005	0.328

(*continued on next page*)

TABLE 3 (Continued)

Panel D: Descriptive Statistics for Firm-Specific Correlations of Individual Fair-Value Income Components over the Period 1996–2004

Pearson Correlations of	Negative Count Signif./Total	Positive Count Signif./Total	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
RG(L) with UG(L) on available-for-sale securities	8/120	3/82	−0.078	0.325	−0.300	−0.058	0.120
UG(L) on available-for-sale securities with CI	0/6	123/196	0.680	0.285	0.524	0.753	0.907
CI with UG(L) on financial assets	3/48	15/154	0.249	0.336	0.020	0.287	0.511
CI with UG(L) on financial liabilities	14/138	1/64	−0.187	0.359	−0.470	−0.211	0.032
UG(L) on financial assets with UG(L) on financial liabilities	65/164	2/38	−0.398	0.401	−0.710	−0.487	−0.139
CI with UG(L) on non-term deposits	30/181	0/21	−0.373	0.288	−0.584	−0.419	−0.193

σ_{NI} denotes firm-specific standard deviation in net income scaled as a percentage of average total assets each year over 1996–2004.

σ_{CI} denotes firm-specific standard deviation in comprehensive income scaled as a percentage of average total assets each year over 1996–2004.

σ_{FFVI} denotes firm-specific standard deviation in full fair-value income scaled as a percentage of average total assets each year over 1996–2004.

Significance level in Panels B and C is $p < 0.05$.

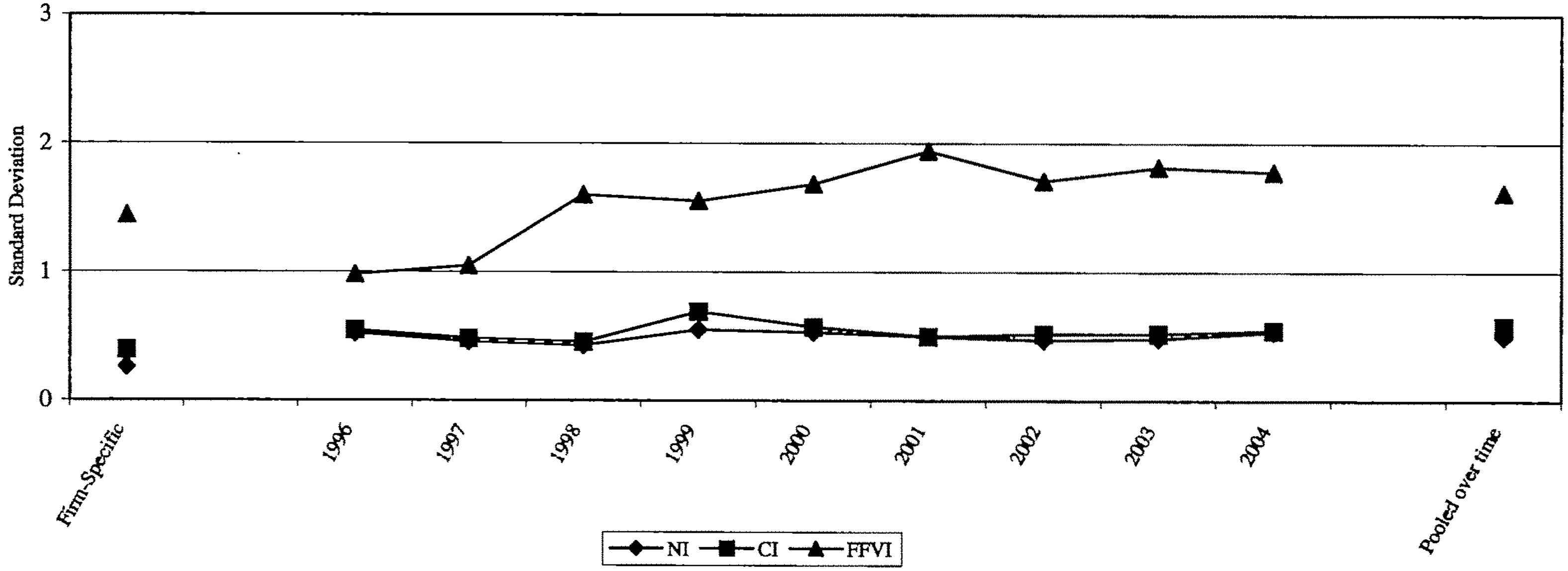
RG(L) denotes realized gain (loss)

UG(L) denotes unrealized gain (loss)

RG(L) and UG(L) on available-for-sale securities, financial assets, financial liabilities, and non-term deposits are scaled as a percentage of average total assets in each year.

n = 202 sample banks.

FIGURE 2
Standard Deviations of Alternate Income Measures, in Cross-Section and over Time



The data points in this figure reflect standard deviations of each income measure (net income (NI), comprehensive income (CI), and full fair-value income (FFVI)) as a percentage of average total assets each year, computed over the 1996–2004 sample period. The data points in the left-most column of the figure represent the means of the firm-specific standard deviations. The yearly and pooled data points represent cross-sectional average standard deviations in each sample year and pooled over all sample years.
n = 202 sample banks.

194 banks (96.0 percent), with 155 banks (76.7 percent) experiencing significantly greater volatility in FFV income (median F-statistic = 7.6).

For FFV-income volatility to be lower than comprehensive income volatility, the variance of the FFV adjustments would have to be less than twice the negative covariance between comprehensive income and the FFV adjustments.[11] Our analysis of the covariance structure among components of FFV income, summarized in Table 3, Panels C and D, suggests that this is not the case. Table 3, Panel C reports descriptive statistics of the firm-specific correlations among the income adjustments necessary to adjust from net income to comprehensive income to FFV income. These correlations indicate that for the average sample bank, comprehensive income adjustments are uncorrelated with net income. The correlations also show that for the average sample bank, FFV adjustments are uncorrelated with comprehensive income and with comprehensive income adjustments. For 15 sample banks, however, FFV adjustments are significantly negatively correlated with comprehensive income adjustments implying that for these banks FFV adjustments capture a significant degree of hedging of the fair value gains and losses that impact comprehensive income. However, for seven other banks FFV adjustments are significantly positively correlated with comprehensive income adjustments implying FFV adjustments reveal that these banks are exposed to a significant degree of additional risk that is not captured in the fair value gains and losses that affect comprehensive income.

To further explore how various fair-value income components covary, we report in Table 3, Panel D descriptive statistics on firm-specific correlations among fair-value gains and losses of individual components of income that comprise the necessary adjustments to compute comprehensive income and FFV income. These statistics indicate that realized and unrealized gains and losses on available-for-sale investment securities for the average bank are essentially uncorrelated. Not surprisingly, unrealized gains/losses on available-for-sale investment securities (a component of comprehensive income) correlate positively with comprehensive income for the majority of sample banks. Comprehensive income also covaries positively with unrealized gains/losses on financial assets that are not recognized at fair value (e.g., held-to-maturity securities and loans) for 154 sample banks (76 percent of the sample; 15 significant), and negatively with unrealized gains/losses on financial liabilities (e.g., term deposits and long-term financial liabilities) for 138 sample banks (68 percent of the sample; 14 significant). Moreover, 164 of 202 sample banks (81 percent) experience negative correlation between unrealized gains/losses on financial assets and financial liabilities, including 65 banks for which it is significantly negative. Finally, these descriptive statistics suggest that unrealized gains/losses on non-term deposits correlate negatively with comprehensive income for 181 sample banks (90 percent of the sample; 30 significant). Together, these descriptive statistics on correlations among fair-value income components indicate that volatility in comprehensive income does not fully capture the complex covariation among the various financial assets and liabilities that comprise bank balance sheets and risk-management strategies.

---

[11] Because comprehensive income (CI) is equal to net income (NI) plus other comprehensive income items (OCI) the variance in comprehensive income can be expressed as $\sigma^2_{CI} = \sigma^2_{NI} + \sigma^2_{OCI} + 2\text{Cov}(NI, OCI)$. In similar fashion, our measure of FFV income is equal to comprehensive income plus all other fair-value adjustments (FVA), so the variance in FFV income (FFVI) can be expressed as: $\sigma^2_{FFVI} = \sigma^2_{CI} + \sigma^2_{FVA} + 2\text{Cov}(CI, FVA)$. The extent to which σ^2_{FFVI} differs from σ^2_{CI} depends on the relative magnitude of σ^2_{FVA} and the sign and magnitude of Cov(CI, FVA). The Cov(CI, FVA) term represents the covariance between comprehensive income and all other FVA items (i.e., fair-value gains and losses on all other financial instruments).

Tests of the Relations between Income Volatility Measures and Market-Based Risk Measures

We assess whether our income-volatility measures capture market-based risk factors by examining their associations with stock-market-based risk measures and banks' disclosed measures of exposure to market risks. We estimate the correlations between each of the three income-volatility measures and four measures of banks' stock-market-based risk: market model beta, total volatility in stock returns, and the absolute values of short-term and long-term interest-rate betas. Using data from banks' market risk disclosures, we also examine whether our income-volatility measures correlate with two proxies for banks' exposure to market risks: total exposure to derivatives and the absolute value of the excess of fixed-rate assets over fixed-rate liabilities subject to repricing within one year. If the income volatility measures capture elements of market-based risks, then the correlations between these risk measures and our income volatility measures should be positive. Further, income volatility measures that are more complete measures of banks' market risks should correlate more positively with market-risk measures.

For this analysis and most of the remaining analysis of the paper, we estimate our three income volatility measures using standard deviations of each income measure stated as a percentage of average total assets over five rolling five-year periods, each ending with years 2000–2004 (rather than estimating standard deviations over the full nine-year period 1996–2004 as in Table 3). We adopt a five-year measurement period to allow income volatility and market-based risk factors to vary over time, relaxing the restriction that income volatility and market-based risks are stationary for each bank over nine years.[12] By relaxing this restriction, however, we may be adding noise to our tests because standard deviations of income over five years may contain more measurement error than standard deviations measured over nine years.

We compute for each bank a market-model beta as a proxy for systematic risk under the CAPM. We estimate market model beta by regressing bank stock returns on a market-wide index of value-weighted returns over each of the five 60-month periods ending with the last month of each year 2000–2004. As a proxy for total risk, we compute the standard deviation of raw returns over the same five 60-month periods, each ending with the last month of the year 2000–2004.

Interest-rate risk is a key element of risk for banks, and it relates directly to fair-value gains and losses. We estimate two firm-specific interest-rate risk parameters by regressing bank stock returns on monthly changes in short- and long-term U.S. Treasury yields over each of the five 60-month periods ending with the last month of each year 2000–2004 (Flannery and James 1984). If a bank has net fixed-rate asset exposure[13] (net fixed-rate liability exposure) to interest-rate changes, then the bank's share returns should covary negatively (positively) with changes in interest rates, leading to more negative (more positive) interest-rate betas. Thus, either greater net fixed-rate asset or net fixed-rate liability exposure could trigger greater volatility in income. Therefore, we test the correlations between the three income-volatility measures and the absolute values of short-term and long-term interest-rate betas.

We also examine the relations between our income volatility measures and two disclosed measures of banks' exposure to market risk (GAP and EXP), measured at the end of each year 2000–2004. GAP denotes the absolute value of the excess of fixed-rate assets over fixed-rate liabilities subject to repricing within one year, all scaled by total assets. The maturity composition of banks' assets and liabilities is associated with the sensitivity of their stock prices to interest rate changes (Flannery and James 1984; Schrand 1997). Because GAP risk manifests as change in net interest income, we expect this measure to be most closely associated with net-income volatility. We obtain GAP from SNL DataSource. SNL obtains GAP data from banks' Forms 10-K and if GAP is missing from the Form 10-K, SNL replaces it with one-year regulatory GAP obtained from banks' year-end Y9-C regulatory reports.

---

[12] Our inferences remain unchanged if we assume that income volatility and market risks are stationary and conduct our analyses on firm-specific standard deviations of income during the full nine-year sample period.

[13] Net fixed-rate asset exposure arises when a bank has a greater proportion of fixed-rate assets than fixed-rate liabilities.

The second disclosed measure of banks' exposure to market risk is the total notional value of derivatives held at year-end, which we denote EXP. Prior research has documented an association between derivatives use and bank stock prices; however, the direction of association is unclear. For example, Riffe (1996) finds that the notional amount of swaps is positively associated with equity values. Riffe (1996) interprets the positive coefficient as evidence that the notional value of swaps is a proxy for higher expected fee income. In contrast, Venkatachalam (1996) finds that the notional value of derivatives is negatively related to stock prices, after controlling for the fair values of derivatives. Venkatachalam (1996) speculates that notional values provide information about risk that may not be fully reflected in derivative fair value estimates. Neither of these studies includes income or risk measures in tests of association between derivatives disclosures and equity values. In our study, EXP equals the sum of the notional values of futures contracts, forwards contracts, exchange-written and exchange-purchased options, OTC-written and OTC-purchased options, and swaps, all scaled by total assets, at the end of each year 2000–2004. We obtain EXP from SNL DataSource, and SNL obtains these data from Y9-C regulatory reports.

Table 4, Panel A contains descriptive statistics for the four stock-market-based risk measures and the two disclosed market-risk measures for our sample of banks for the years 2000–2004. The mean (median) market model beta is 0.312 (0.299), well below the market-wide average of 1.0. Mean (median) absolute values of short-term and long-term interest-rate betas are approximately 0.141 (0.101) and 0.225 (0.163), respectively. The mean (median) standard deviation in raw returns for our sample banks is 0.051 (0.048). The mean (median) repricing GAP in absolute value is roughly 12.6 percent (10.4 percent) of total assets. More than half of our sample banks have no exposure to derivatives (EXP), whereas the mean bank holds derivatives that amount to roughly 37.3 percent of total assets. This skewed distribution is not surprising, given prior findings (e.g., Venkatachalam 1996) that larger, more sophisticated banks are more actively involved in derivatives trading than are smaller banks.[14]

Table 4, Panel B contains Pearson and Spearman correlation statistics among the three income-volatility measures, the four stock-market-based risk measures, and the two disclosed measures of market risk, pooled across years 2000–2004. The most consistent set of results emerge from the correlations between the three income-volatility measures and the standard deviation of returns, all of which are positive and significant. These results are intuitive in that volatility in reported income and volatility in returns are general measures of total risk, and are not specific to certain types of risk. These findings confirm that our three income-volatility measures capture risk factors that relate to volatility in stock returns.

---

[14] In untabulated analyses, we find similar descriptive statistics for our four stock-market-based risk measures and our two disclosed measures of market risk for the full sample of banks with at least one year of complete accounting and stock-market-based variables available during 2000–2004. These descriptive statistics suggest our final sample banks have market-risk profiles comparable to the larger population of banks.

TABLE 4
Associations between Bank Income Volatility Measures, Market-Based Risk Measures, and Disclosed Measures of Market Risks

Panel A: Market-based risk measures and disclosed market-risk measures across sample banks with complete financial statement information for the period 1996–2004.

	Mean	Std. Dev.	Min	1st Quartile	Median	3rd Quartile	Max
Market model beta (MMBeta)	0.312	0.275	−0.609	0.137	0.299	0.464	1.682
Short-term interest-rate beta (STBeta)	0.141	0.142	0.000	0.047	0.101	0.187	1.128
Long-term interest-rate beta (LTBeta)	0.225	0.258	0.000	0.067	0.163	0.299	1.979
Standard deviation of raw returns (σRet)	0.051	0.016	0.022	0.042	0.048	0.056	0.180
Interest-rate repricing gap (GAP)	0.126	0.106	0.000	0.044	0.104	0.181	0.593
Derivatives exposure (EXP)	0.373	2.769	0.000	0.000	0.000	0.045	43.016

Panel B: Pearson (upper triangle) and Spearman (lower triangle) Correlations: Among and between standard deviations of income metrics, market-based risk metrics, and disclosed measures of market-risks across banks with complete financial statement information, for each of the five-year periods ending 2000–2004.

	σNI	σCI	σFFVI	σRet	MMBeta	STBeta	LTBeta	GAP	EXP
σNI	1.000	0.717***	0.133***	0.475***	−0.018	0.172***	0.163***	0.082***	0.054*
σCI	0.420***	1.000	0.121***	0.354***	−0.109***	0.104***	0.097***	0.051	−0.038
σFFVI	0.187***	0.182***	1.000	0.408***	0.077**	0.060*	0.604***	0.041	0.104***
σRet	0.349***	0.128***	0.172***	1.000	0.127***	0.330***	0.504***	0.049	0.124***
MMBeta	0.090***	−0.030	0.013	0.322***	1.000	0.171***	0.165***	−0.032	0.375***
STBeta	0.135***	0.107***	0.051	0.305***	0.141***	1.000	0.246***	0.013	0.063**
LTBeta	0.076**	−0.025	0.056	0.288***	0.184***	0.170***	1.000	0.014	0.070**
GAP	0.034	0.004	0.056*	0.009	−0.008	−0.008	−0.025	1.000	0.002
EXP	0.095***	0.061*	0.139***	0.158***	0.407***	0.031	0.194***	0.022	1.000

(*continued on next page*)

TABLE 4 (Continued)

Panel C: Parameter Estimates from Regressions of Banks' Market-Based Risk Measures on Income Volatility Measures and Disclosed Market-Risk Measures, All Sample Banks across 2000–2004 (n = 1,010; t-statistics in parentheses)

Dependent Variable	Model	EXP	GAP	σ_{NI}	σ_{CI}	σ_{FFVI}	$\sigma_{CI} - \sigma_{NI}$	$\sigma_{FFVI} - \sigma_{CI}$	R^2
σRet	(1)	0.0006	0.0028	0.0374					0.262
		(3.69)***	(0.67)	(17.28)***					
	(2)	0.0008	0.0057		0.0217				0.165
		(4.82)***	(1.29)		(12.14)***				
	(3)	0.0005	0.0061			0.0026			0.205
		(2.97)***	(1.44)			(14.31)***			
	(4)	0.0004	0.0010	0.0351			0.0003	0.0023	0.383
		(2.64)***	(0.28)	(17.42)***			(0.15)	(14.04)***	
MMBeta	(1)	0.0381	−0.0741	−0.0459					0.204
		(13.60)***	(−1.01)	(−1.20)					
	(2)	0.0376	−0.0691		−0.0884				0.210
		(13.48)***	(−0.95)		(−2.98)***				
	(3)	0.0373	−0.0875			0.0065			0.207
		(13.29)***	(−1.20)			(2.08)**			
	(4)	0.0370	−0.0781	−0.0327			−0.1127	0.0070	0.213
		(13.19)***	(−1.07)	(−0.84)			(−2.83)***	(2.26)**	
LTBeta	(1)	0.0059	0.0139	0.2027					0.046
		(2.07)**	(0.19)	(5.20)***					
	(2)	0.0070	0.0331		0.0920				0.029
		(2.45)**	(0.44)		(3.01)***				
	(3)	0.0009	−0.0145			0.0629			0.394
		(0.39)	(−0.24)			(24.88)***			
	(4)	0.0002	−0.0326	0.1524			−0.0544	0.0617	0.406
		(0.10)	(−0.56)	(4.86)***			(−1.69)	(24.68)***	

(continued on next page)

TABLE 4 (Continued)

---

*, **, *** Denotes $p < .10$, $p < .05$, and $p < .01$, respectively, all two-tailed.

σ_{NI}, σ_{CI}, and σ_{FFVI} denote firm-specific standard deviation in annual net income, comprehensive income, and full fair-value income, respectively, expressed as a percentage of average total assets; we measures these standard deviations over each of the five year periods ending with year *t* (2000–2004).

σ_{RET} denotes firm-specific standard deviation in monthly stock returns, which we measure over each of the five year periods ending with year *t* (2000–2004).

GAP denotes the absolute value of the excess of interest-bearing assets over interest-bearing liabilities subject to repricing within one year measured at the end of each year *t* (2000–2004). We obtain GAP from banks' Forms 10-K filed with the SEC and scale it by total assets. If GAP is missing from the Form 10-K, then we replace it with one-year regulatory GAP obtained from banks' year-end Y9-C regulatory reports.

EXP denotes total exposure arising from derivatives held at the end of each year *t* (2000–2004), scaled by total assets. We obtain EXP from Y9-C regulatory reports, and it equals the sum of the notional value of futures contracts, forwards contracts, exchange-written and exchange-purchased options, OTC-written and OTC-purchased options, and swaps.

MMBeta denotes the absolute-value of market model beta estimated using bank stock returns and a value-weighted index of returns, measured over the 60-month period ending with the last month of year *t* (2000–2004).

STBeta and LTBeta denote the absolute-value of short-term and long-term interest-rate beta, respectively, measured over the 60-month period ending with the last month of year *t* (2000–2004).

Panel C reports regression analyses of three market-based risk proxies, σ_{RET}, MMBeta, and LTBeta (each denoted MktRiskProxy in the general models below) on the income volatility metrics and disclosed market-risk proxies (EXP and GAP) for each of the five years 2000–2004. Intercepts are not tabulated.

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\, \delta_{2000\text{–}2003} + \delta_1 \text{EXP}_{jt} + \delta_2 \text{GAP}_{jt} + \delta_3 \sigma_{NIjt} + \varepsilon_{jt} \tag{1}$$

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\, \delta_{2000\text{–}2003} + \delta_1 \text{EXP}_{jt} + \delta_2 \text{GAP}_{jt} + \delta_3 \sigma_{CIjt} + \varepsilon_{jt} \tag{2}$$

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\, \delta_{2000\text{–}2003} + \delta_1 \text{EXP}_{jt} + \delta_2 \text{GAP}_{jt} + \delta_3 \sigma_{FFVIjt} + \varepsilon_{jt} \tag{3}$$

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\, \delta_{2000\text{–}2003} + \delta_1 \text{EXP}_{jt} + \delta_2 \text{GAP}_{jt} + \delta_3 \sigma_{NIjt} + \delta_4(\sigma_{CIjt} - \sigma_{NIjt}) + \delta_5(\sigma_{FFVIjt} - \sigma_{CIjt}) + \varepsilon_{jt} \tag{4}$$

The results in Table 4, Panel B do not support reliable inferences about whether any of the three income-volatility measures correlate with systematic risk (MMBeta). Although net income volatility and FFV-income volatility exhibit some evidence of positive correlation with MMBeta, none of the three income volatility measures exhibits significant positive coefficients for both Pearson and Spearman correlations.[15] The correlation coefficients between the three income volatility measures and the interest rate risk measures suggest income volatility covaries positively with interest rate risk. In particular, net income volatility is consistently positively correlated with both STBeta and LTBeta, and comprehensive income volatility is consistently positively correlated with STBeta. The correlations between FFV-income volatility and STBeta and LTBeta are mixed. Both Pearson correlation coefficients are positive and significant (especially with LTBeta) but both Spearman correlation coefficients are insignificant. The results are also weak for correlations with net asset-liability-repricing risk (GAP). None of the three income-volatility measures has consistently significant positive Pearson and Spearman correlation coefficients with GAP. With regard to exposure to derivatives (EXP), net-income volatility and FFV-income volatility exhibit consistent positive and significant correlations with EXP. Summarizing, these correlations indicate that, among our three income-volatility measures, volatility in net income emerges as the most robust and consistent measure of risk across the set of risk factors examined here. Comprehensive-income volatility appears to reflect a subset of risks, most notably volatility in stock returns and short-term interest rate risk. FFV-income volatility reflects elements of risk captured by the standard deviation in stock returns and derivatives exposure.

We also explore the incremental market-risk-related information content in our three income-volatility measures with two sets of tests.[16] First, we examine whether our three income-volatility measures provide any market-risk-relevant information beyond the two market-risk measures that banks already disclose (i.e., GAP and EXP). Second, we examine whether incremental volatility in comprehensive income and FFV income provide market-risk-relevant information beyond the volatility in net income, and controls for GAP and EXP. To run these tests, we focus on three market risk proxies: MMBeta as a proxy for systematic risk, standard deviation in stock returns as a proxy for total risk, and LTBeta as a proxy for interest rate risk (each denoted MktRiskProxy in the general models below). We estimate the following regression models pooled over the five years 2000–2004:

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\delta_{2000\text{–}2003} + \delta_1\text{EXP}_{jt} + \delta_2\text{GAP}_{jt} + \delta_3\sigma_{\text{NI}jt} + \varepsilon_{jt}; \tag{1}$$

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\delta_{2000\text{–}2003} + \delta_1\text{EXP}_{jt} + \delta_2\text{GAP}_{jt} + \delta_3\sigma_{\text{CI}jt} + \varepsilon_{jt}; \tag{2}$$

$$\text{MktRiskProxy}_{jt} = \delta_0 + \Sigma_4\delta_{2000\text{–}2003} + \delta_1\text{EXP}_{jt} + \delta_2\text{GAP}_{jt} + \delta_3\sigma_{\text{FFVI}jt} + \varepsilon_{jt}; \tag{3}$$

$$\begin{aligned}\text{MktRiskProxy}_{jt} = {} & \delta_0 + \Sigma_4\delta_{2000\text{–}2003} + \delta_1\text{EXP}_{jt} + \delta_2\text{GAP}_{jt} + \delta_3\sigma_{\text{NI}jt} \\ & + \delta_4(\sigma_{\text{CI}jt} - \sigma_{\text{NI}jt}) + \delta_5(\sigma_{\text{FFVI}jt} - \sigma_{\text{CI}jt}) + \varepsilon_{jt}.\end{aligned} \tag{4}$$

Models (1), (2), and (3) test whether any of our three income volatility measures provide incremental risk-relevant information beyond GAP and EXP. If so, then δ_3 should be positive and significant. Model (4) tests whether incremental volatility in comprehensive income and/or FFV income provides any market-risk-relevant information beyond the volatility in net income, and controls for GAP and EXP. If so, then δ_4 and/or δ_5 should be positive.

---

[15] We do not find consistent evidence of significant positive association between net income volatility and market-model beta, inconsistent with prior research on non-bank samples (e.g., Beaver et al. 1970).

[16] We gratefully acknowledge a reviewer for suggesting this analysis.

Table 4, Panel C reports the parameter estimates of these models. For the standard deviation in stock returns, the results for models (1), (2), and (3) indicate that all three income-volatility measures provide incremental risk-relevant information. The results for model (4) indicate that incremental volatility in FFV income provides significant incremental risk-relevant information for total risk in returns, beyond net-income volatility, GAP and EXP. The results from all four models for MMBeta indicate that only volatility in FFV income provides incremental information for systematic risk. Finally, all four models indicate that all three income-volatility measures provide incremental information for LTBeta, beyond GAP and EXP. The findings for model (4) suggest that volatility in FFV income provides the most significant incremental information for LTBeta, as a proxy for interest rate risk.

Tests of Association between Income Volatility Measures and Bank Share Prices

To address our third and final research question, we conduct two related sets of tests of association between volatilities of the three income measures and share prices. Both sets of tests provide evidence on our prediction that income volatility is an element of risk that increases expected returns and decreases share prices, *ceteris paribus*. First, we assess the share-price implications of each income-volatility measure using a regression approach motivated by the residual-income model (Ohlson 1995). Second, we test the implications of each income-volatility measure for expected rates of return implied by bank share prices.

Share-Price Tests Utilizing a Residual Income Approach

Model (5) is a highly simplified version of the residual-income model (Ohlson 1995). We estimate Equation (5) as a benchmark model before introducing any income-volatility measures, on a pooled sample over 2000–2004:

$$P_{jt} = \alpha_0 + \Sigma_4\alpha_{2000-2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \varepsilon_{jt}. \tag{5}$$

In this equation, P denotes price per share for bank j at the end of t; BVE denotes book value of equity per share at the end of year t; and AE denotes period t abnormal earnings per share, which we use as a proxy for expected future abnormal earnings. We measure abnormal earnings as current period earnings (scaled by the number of shares outstanding) less the product of the risk-free rate of return at the beginning of year t times book value per share at the beginning of year t. We use the risk-free rate for this computation to allow our coefficient estimates to capture the effects of risk. To control for sample-wide mean shifts in share prices over our study period, we allow the intercept to shift each year (α_{2000}–α_{2003}).[17]

Theoretically, the coefficient estimate on BVE (α_1) should equal 1, but omitted variables (e.g., off-balance-sheet net asset values and gains or losses that have not yet been realized or recognized) may induce this coefficient estimate to deviate from 1. We expect the coefficient on AE (α_2) to be positive, reflecting market capitalization of abnormal earnings. Because this simplified model only includes abnormal earnings for the current period (rather than all future periods), this coefficient should reflect the cross-sectional mean share-price effects of persistence, risk, growth, and other information in abnormal earnings (consistent with the linear information dynamics assumptions in Ohlson [1995]).[18]

---

[17] The α_0 coefficient estimate can therefore be interpreted as the intercept for the year 2004, the fifth and final year of the sample in this pooled regression.

To control for information in banks' disclosed measures of exposure to interest-rate and market risks, we augment Equation (5) with GAP and EXP. In the model, we interact GAP with AE and predict that the coefficient will be negative, as the market assigns a lower capitalization multiple to abnormal earnings of banks with greater exposure to interest-rate risk (i.e., larger GAP), *ceteris paribus*.

As described above, prior research suggests that EXP, our proxy for banks' exposure to derivatives, may be associated with higher levels of profitability (Riffe 1996), greater risk (Venkatachalam 1996), and may be a proxy for larger banks that are more actively involved in derivatives (Venkatachalam 1996). To distinguish the confounding effects of higher risk, higher returns, and size associated with derivatives on share prices, we include in the regression a main effect for EXP as well as the interaction between EXP and AE. We interpret a positive coefficient on EXP as consistent with EXP acting as a proxy for larger banks with net unrecorded assets associated with derivatives. We expect the coefficient on EXP interacted with AE to be negative because we predict that the capital markets assign lower valuation multiples to abnormal earnings for banks with greater exposure to derivatives (i.e., greater derivatives exposure is more risky for banks that are less profitable). With these controls for banks' exposure to interest-rate and market risks, we estimate the following regression:

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4\alpha_{2000\text{–}2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) \\ & + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \varepsilon_{jt}. \end{aligned} \tag{6}$$

We test the share-price effects of the volatilities of the three income metrics by separately interacting each with abnormal earnings, thereby allowing the valuation coefficient on abnormal earnings to vary across each income-volatility measure. We estimate the following models:

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4\alpha_{2000\text{–}2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) \\ & + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) + \varepsilon_{jt}; \end{aligned} \tag{7}$$

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4\alpha_{2000\text{–}2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) \\ & + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{CIjt} \times AE_{jt}) + \varepsilon_{jt}; \end{aligned} \tag{8}$$

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4\alpha_{2000\text{–}2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) \\ & + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{FFVIjt} \times AE_{jt}) + \varepsilon_{jt}. \end{aligned} \tag{9}$$

For these models, we compute our income-volatility variables (i.e., σ_{NIjt}, σ_{CIjt}, and σ_{FFVIjt}) as bank-specific standard deviations of each income measure expressed as a percentage of average total assets. We measure each observation on a rolling basis over the five-year periods ending with each year t (2000–2004). In each model, a significant negative α_6 coefficient suggests that the income-volatility metric captures elements of risk that are priced in the capital markets, after controlling for GAP and EXP.

---

[18] In 58 bank-years, abnormal earnings are negative. To avoid the unrealistic assumption of negative abnormal earnings in perpetuity, we set abnormal earnings for those bank-years equal to that bank's time-series mean abnormal earnings. For the few banks with negative time-series mean abnormal earnings, we set negative abnormal earnings for year *t* equal to 0.01.

To test the share-price effects of the incremental volatilities of σ_{CI} and σ_{FFVI} beyond σ_{NI} we run the following regression:

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4 \alpha_{2000\text{–}2003} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) \\ & + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) \\ & + \alpha_7[(\sigma_{CIjt} - \sigma_{NIjt}) \times AE_{jt}] + \alpha_8[(\sigma_{FFVIjt} - \sigma_{CIjt}) \times AE_{jt}] + \varepsilon_{jt}. \end{aligned} \tag{10}$$

Equation (10) tests whether the capital markets price as risk the incremental volatility associated with CI ($\sigma_{CI} - \sigma_{NI}$) and FFVI ($\sigma_{FFVI} - \sigma_{CI}$).[19] Closely related to Equation (10), we also estimate the following model using bank share prices at the end of year 2004 (i.e., t = 2004) and using income-volatility metrics estimated over the full nine-year sample period:

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) \\ & + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) + \alpha_7[(\sigma_{CIjt} - \sigma_{NIjt}) \times AE_{jt}] \\ & + \alpha_8[(\sigma_{FFVIjt} - \sigma_{CIjt}) \times AE_{jt}] + \varepsilon_{jt}. \end{aligned} \tag{11}$$

For both Equations (10) and (11), we interpret significant negative α_7 and α_8 coefficient estimates as evidence that these incremental income-volatility metrics represent elements of risk priced by the capital markets. As noted earlier, we predict σ_{FFVI} should be the most complete measure of income volatility, and the results in Table 4, Panel C indicate that σ_{FFVI} contains significant incremental information for the standard deviation in stock returns, market-model beta, and long-term interest-rate beta. Thus, we predict that incremental volatility in σ_{FFVI} will have a significant negative association with the capitalization of abnormal earnings in bank-share prices.

Table 5 reports the results of estimating models (5) through (11), as well as descriptive statistics on variables used only in these regressions (i.e., P, BVE, AE, and AE interacted with the volatility metrics). The coefficient estimates on BVE are consistently positive and significant, but they are statistically less than the theoretically predicted value of 1 across all models except for model (11). The coefficients on AE are consistently positive and significant, ranging from 10.19 to 12.60 across models. When interacted with AE, GAP is negatively associated with bank share prices, consistent with our predictions that share prices reflect lower capitalization multiples to abnormal earnings for banks with greater exposure to interest-rate risk. The results also indicate that across models (with the exception of model (11)), EXP is consistently positively associated with bank share prices, and when interacted with AE, it is consistently negatively associated with share prices.

With regard to our main predictions, results for models (7) and (9) indicate the coefficients on the NI and FFVI volatility metrics are negative and significant. In particular, the coefficient on the interaction of σ_{FFVI} × AE in model (9) is negative and significant (−0.21; t-statistic = −3.54), consistent with σ_{FFVI} capturing risk-relevant variation in valuation of

---

[19] This decomposition of standard deviation is consistent with tests run in prior studies (e.g., Barth et al. 1995) and our discussion of the incremental effect of each income-volatility coefficient assumes that variances are additive.

TABLE 5
Association between Stock Price, Book Value, Abnormal Earnings, Risk Proxies, and Income Volatility

Models:

$$P_{jt} = \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \varepsilon_{jt} \quad (5)$$

$$P_{jt} = \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \varepsilon_{jt} \quad (6)$$

$$P_{jt} = \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) + \varepsilon_{jt} \quad (7)$$

$$P_{jt} = \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{CIjt} \times AE_{jt}) + \varepsilon_{jt} \quad (8)$$

$$P_{jt} = \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{FFVIjt} \times AE_{jt}) + \varepsilon_{jt} \quad (9)$$

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \Sigma_4\, \alpha_{2000\text{–}2004} + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) \\ & + \alpha_7[(\sigma_{CIjt} - \sigma_{NIjt}) \times AE_{jt}] + \alpha_8[(\sigma_{FFVIjt} - \sigma_{CIjt}) \times AE_{jt}] + \varepsilon_{jt} \end{aligned} \quad (10)$$

$$\begin{aligned} P_{jt} = {} & \alpha_0 + \alpha_1 BVE_{jt} + \alpha_2 AE_{jt} + \alpha_3(GAP_{jt} \times AE_{jt}) + \alpha_4(EXP_{jt} \times AE_{jt}) + \alpha_5 EXP_{jt} + \alpha_6(\sigma_{NIjt} \times AE_{jt}) + \alpha_7[(\sigma_{CIjt} - \sigma_{NIjt}) \times AE_{jt}] \\ & + \alpha_8[(\sigma_{FFVIjt} - \sigma_{CIjt}) \times AE_{jt}] + \varepsilon_{jt} \end{aligned} \quad (11)$$

(continued on next page)

TABLE 5 (Continued)

Parameter Estimates: (Models 5–10, n = 1,010; Model 11, n = 202; t-statistics in parentheses)

	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Intercept	8.20 (11.45)***	8.21 (11.48)***	8.31 (11.61)***	8.26 (11.53)***	8.23 (11.57)***	8.32 (11.69)***	3.67 3.77***
2000	−4.24 (−5.29)***	−4.23 (−5.28)***	−4.23 (−5.28)***	−4.19 (−5.21)***	−4.25 (−5.34)***	−4.42 (−5.53)***	
2001	−3.44 (−4.34)***	−3.49 (−4.41)***	−3.51 (−4.44)***	−3.45 (−4.35)***	−3.47 (−4.41)***	−3.67 (−4.66)***	
2002	−5.85 (−7.52)***	−5.93 (−7.65)***	−5.93 (−7.66)***	−5.85 (−7.50)***	−5.93 (−7.68)***	−6.08 (−7.85)***	
2003	−1.52 (−1.96)**	−1.50 (−1.95)**	−1.50 (−1.95)**	−1.40 (−1.80)**	−1.49 (−1.94)**	−1.64 (−2.12)**	
BVE	0.77 (23.69)***	0.75 (22.67)***	0.73 (21.63)***	0.74 (21.44)***	0.74 (22.36)***	0.74 (21.15)***	1.00 18.24***
AE	10.19 (25.66)***	11.02 (21.57)***	11.47 (20.59)***	11.38 (18.43)***	11.50 (22.06)***	11.33 (17.07)***	12.60 11.29***
GAP × AE		−0.04 (−2.21)**	−0.03 (−1.88)*	−0.03 (−2.02)**	−0.04 (−2.23)**	−0.03 (−1.98)**	−0.09 −3.37***
EXP × AE		−0.21 (−2.06)**	−0.20 (−1.95)**	−0.21 (−2.09)**	−0.19 (−1.84)*	−0.17 (−1.69)*	−0.17 −0.68
EXP		0.35 (2.21)**	0.36 (2.22)**	0.35 (2.24)**	0.36 (2.25)**	0.35 (2.21)**	0.06 0.24
σ_{NI} × AE			−1.97 (−1.93)**			−1.88 (−1.83)*	−2.80 −1.55
σ_{CI} × AE				−0.97 (−1.02)			
σ_{FFVI} × AE					−0.21 (−3.54)***		
$(\sigma_{CI} - \sigma_{NI})$ × AE						1.93 (1.56)	3.46 1.60
$(\sigma_{FFVI} - \sigma_{CI})$ × AE						−0.17 (−3.07)***	−0.31 −2.47***
Adjusted R^2	0.768	0.770	0.770	0.770	0.773	0.773	0.863

(continued on next page)

TABLE 5 (Continued)

Descriptive Statistics for the Regression Variables (n = 1010)

Variable	Mean	Std. Dev.	Min	1st Quartile	Median	3rd Quartile	Max
P	32.98	18.26	1.81	22.70	28.66	37.50	148.25
BVE	15.67	10.54	0.73	10.33	13.50	17.78	124.05
AE	1.24	0.87	0.01	0.67	1.09	1.50	4.69
$\sigma_{NI} \times AE$	0.32	0.39	0.00	0.11	0.19	0.36	3.16
$\sigma_{CI} \times AE$	0.46	0.42	0.00	0.21	0.35	0.57	2.79
$\sigma_{FFVI} \times AE$	1.90	4.15	0.01	0.55	0.94	1.90	45.79

*, **, *** Denotes $p < .10$, $p < .05$, and $p < .01$, respectively, all two-tailed.

Tables 3 and 4 present descriptive statistics for the regressor variables σ_{NI}, σ_{CI}, σ_{FFVI}, GAP, and EXP, so we do not duplicate them here.

P denotes price per share at the end of year *t*.

BVE denotes book value of equity per share at the end of year *t*.

AE denotes abnormal earnings per share, which we compute as reported earnings per share for year *t* minus the risk-free rate of return at the beginning of year *t* times the book value of equity per share at the beginning of year *t*.

GAP denotes the absolute value of the excess of interest-bearing assets over interest-bearing liabilities subject to repricing within one year. We obtain GAP from banks' Form 10-K filed with the SEC and scale it by total assets. If GAP is missing from the Forms 10-K, then we replace it with one-year regulatory GAP obtained from banks' year-end Y9-C regulatory reports.

EXP denotes total exposure arising from derivatives held at the end of each year *t* (2000–2004), scaled by total assets. We obtain EXP from Y9-C regulatory reports, and it equals the sum of the notional value of futures contracts, forwards contracts, exchange-written and exchange-purchased options, OTC-written and OTC-purchased options, and swaps.

For Models 5–10:

σ_{NI} = firm-specific standard deviation of annual net income expressed as a percentage of average total assets as measured over each five-year period ending with year *t* (t = 2000–2004);

σ_{CI} = firm-specific standard deviation of annual comprehensive income expressed as a percentage of average total assets as measured over each five-year period ending with year *t* (t = 2000–2004); and

σ_{FFVI} = firm-specific standard deviation of annual full fair-value income expressed as a percentage of average total assets as measured over each five-year period ending with year *t* (t = 2000–2004).

For Model 11:

σ_{NI} = firm-specific standard deviation (over the full nine-year sample period, 1996–2004) in annual net income expressed as a percentage of average total assets each year;

σ_{CI} = firm-specific standard deviation (over the full nine-year sample period, 1996–2004) in annual comprehensive income expressed as a percentage of average total assets each year; and

σ_{FFVI} = firm-specific standard deviation (over the full nine-year sample period, 1996–2004) in annual full fair-value income expressed as a percentage of average total assets each year.

residual income. By contrast, the coefficient on $\sigma_{CI} \times$ AE in model (8) is negative but not significant. This suggests that volatility in CI obscures the risk-relevant information represented by volatility in NI and FFVI.

The results from estimating models (10) and (11) suggest that volatility in FFV income captures incremental factors that are associated with the market's assessment of share-price risk, beyond the risk factors represented by volatility in net income and comprehensive income, as well as GAP and EXP. Taken further, the significance of the incremental volatility in σ_{FFVI} suggests that σ_{NI}, σ_{CI}, GAP and EXP are incomplete measures of income volatility and risk exposure for purposes of pricing bank shares.[20]

Tests Utilizing a Cost of Capital Approach

We predict that volatility in income represents elements of risk that increase expected returns and decrease share prices, *ceteris paribus*. To refine our analysis to focus directly on expected returns, we examine whether banks' costs of equity capital are associated with each income-volatility measure. If our income-volatility metrics are proxies for elements of risk that the capital markets price in expected returns, then they should be positively associated with banks' costs of equity capital.

We derive our cost of equity capital measure using the model presented in Ohlson and Juettner-Nauroth (2005) (OJ). Using a dividend valuation model and imposing a number of assumptions about near-term earnings and abnormal earnings, OJ derive the following earnings-based-valuation formula:

$$P_0 = \frac{eps_1}{CoE} + \frac{eps_1\left(\dfrac{eps_2 - eps_1}{eps_1} + \dfrac{CoE^*dps_1}{eps_1} - CoE\right)}{CoE((1 + CoE) - \gamma)}. \tag{12}$$

where:

eps_1 = forecasted earnings per share in year 1;
eps_2 = forecasted earnings per share in year 2;
dps_1 = forecasted dividends per share;
CoE = the implied cost of equity capital;
γ = 1 + the growth rate in abnormal earnings beyond the forecast horizon; and
P_0 = value per share.

Rearranging Equation (12) leads to the following formula for the cost of capital:

$$CoE = A + \sqrt{A^2 + \frac{eps_1}{P_0} * \left(\frac{eps_2 - eps_1}{eps_1} - (\gamma - 1)\right)}. \tag{13}$$

---

[20] Consistent with the analysis reported in Barth et al. (1995), we also tested the pricing of each income-volatility measure using a price-earnings model. The results (not tabulated) are consistent with the results in Table 5. The findings indicate that, by itself, σ_{NI} is negatively related to price-earnings ratios and incrementally negatively related to share prices when we include the volatilities of the two sets of fair-value-income adjustments in the regression. The findings indicate that σ_{CI} by itself is not related to price-earnings ratios and $(\sigma_{CI} - \sigma_{NI})$ is incrementally positively related to price-earnings ratios (opposite the predicted direction) when we include σ_{NI} and/or $(\sigma_{FFVI} - \sigma_{CI})$ in the regression. Also, consistent with the results in Table 5, the coefficient estimates on σ_{FFVI} and $(\sigma_{FFVI} - \sigma_{CI})$ are significantly negative and robust across alternative specifications.

where:

$$A = \frac{1}{2}\left((\gamma - 1) + \frac{dps_1}{P_0}\right).$$

To compute Equation (13), we use consensus I/B/E/S eps forecasts as of the last date for which I/B/E/S compiles forecasts preceding year-end (which they denote as the last statistical date) and share prices as of year-end from CRSP. As in Gode and Mohanram (2003), we assume γ equals $1 + (r_f - 3\%)$.

To test whether banks' equity capital costs are increasing in our income volatility metrics, we run the following regression, again including controls for each bank's disclosed exposure to interest-rate and market-risk using GAP and EXP, respectively:

$$\begin{aligned} \mathrm{CoE}_{jt} = {} & \alpha_0 + \gamma_1 \mathrm{GAP}_{jt} + \gamma_2 \mathrm{EXP}_{jt} + \gamma_3 \sigma_{\mathrm{NIjt}} + \gamma_4(\sigma_{\mathrm{CIjt}} - \sigma_{\mathrm{NIjt}}) \\ & + \gamma_5(\sigma_{\mathrm{FFVIjt}} - \sigma_{\mathrm{CIjt}}) + \varepsilon_{jt}. \end{aligned} \tag{14}$$

We predict that the coefficient on the control variable for GAP will be positive because we predict that the capital markets assign higher expected rates of return when discounting earnings for banks with greater degrees of exposure to interest rate risk. We cannot make an unambiguous prediction for the sign of the coefficient on EXP in this regression. If EXP serves as a proxy for greater profitability and/or larger bank size, then the coefficient estimate may be negative, reflecting larger, more profitable banks facing lower expected returns. If EXP serves as a proxy for exposure to market risks, then the coefficient estimate may be positive reflecting higher expected returns.

Our primary predictions for these tests are that the coefficients on the incremental income-volatility metrics will be positive if the capital markets assign higher expected rates of return when discounting earnings for banks with greater volatility in income. Again, as in our share-price tests, we predict σ_{FFVI} should be the most complete measure of income volatility. Thus, we predict that incremental volatility in FFV income will have a significant positive association with the cost of bank equity capital.

We estimate model (14) several ways. First, we estimate it with each annual cross-section of sample banks with available I/B/E/S data in years 2000 to 2004.[21] For each of these models, we compute CoE based on share prices at the end of each year, and we compute income-volatility metrics over the five-year period ending with year t. Table 6, Columns (a) through (e) present the results for these five annual cross-sectional estimations. We also estimate model (14) pooled over the five-year period 2000–2004, and introduce separate intercepts for each year 2000–2003 (separate annual intercepts are not tabulated). Table 6, Column (f) contains these results. Finally, we estimate model (14) using CoE estimates as of the end of year 2004 and firm-specific income-volatility metrics computed using nine years of annual income information for our sample from 1996–2004. We present those results in Column (g) of Table 6.

The results in Table 6 indicate that both GAP and EXP are positively associated with the cost of equity capital in 2003 and 2004, and in the five-year pooled estimation. Similarly, the results indicate that σ_{NIj} is positively associated with banks' costs of equity capital in

---

[21] As noted above, we cannot estimate these models over our full sample of banks because not all sample banks are followed by analysts that report their forecasts to I/B/E/S. The subsample varies with forecast data availability each year, with the largest sample including 131 banks in 2004.

TABLE 6
Association between Cost of Equity Capital, Income Volatility, and Disclosed Measures of Market Risks

Model:

$$CoE_{jt} = \alpha_0 + \gamma_1\, GAP_{jt} + \gamma_2\, EXP_{jt} + \gamma_3\, \sigma_{NIjt} + \gamma_4\, (\sigma_{CIjt} - \sigma_{NIjt}) + \gamma_5\, (\sigma_{FFVIjt} - \sigma_{CIjt}) + \varepsilon_{jt} \quad (14)$$

Parameter Estimates (t-statistics in parentheses)

Column	(a)	(b)	(c)	(d)	(e)	(f)	(g)
CoE Estimation as of Year-End	**2000**	**2001**	**2002**	**2003**	**2004**	**2000–2004**	**2004**
Intercept	0.0790 (12.38)***	0.0826 (20.53***)	0.0910 (25.39***)	0.0765 (25.18)**	0.0776 (32.76)***	0.0796 (41.78)***	0.0758 (29.86)***
2000						0.0048 (2.32)**	
2001						0.0005 (0.25)	
2002						0.0085 (4.47)***	
2003						−0.0008 (−0.39)	
GAP	0.0406 (1.55)	0.0059 (0.36)	0.0118 (0.80)	0.0268 (2.38)**	0.0174 (2.02)**	0.0195 (3.04)***	0.0185 (2.22)**
EXP	−0.0016 (−0.75)	0.0001 (0.16)	0.0004 (1.15)	0.0005 (1.84)*	0.0008 (3.31)***	0.0005 (2.70)***	0.0008 (3.23)***
σ_{NI}	−0.0073 (−0.46)	0.0064 (0.60)	0.0097 (1.25)	0.0139 (1.76)*	0.0128 (2.59)**	0.0089 (2.30)**	0.0179 (3.21)***
$\sigma_{CI} - \sigma_{NI}$	0.0043 (0.35)	−0.0081 (−1.01)	−0.0112 (−1.55)	−0.0032 (−0.56)	0.0019 (0.26)	−0.0048 (−1.36)	0.0003 (0.05)
$\sigma_{FFVI} - \sigma_{CI}$	0.0053 (1.77)*	0.0012 (2.35)**	0.0005 (1.67)*	0.0006 (2.05)**	0.0004 (1.59)	0.0007 (3.71)***	0.0007 (1.99)**
n	92	114	129	122	131	588	131
Adjusted R^2	0.015	0.028	0.051	0.151	0.157	0.110	0.189

(continued on next page)

TABLE 6 (Continued)

Descriptive Statistics for the Regression Variables for the Subset of Banks with Available I/B/E/S Forecasts in 2000–2004 (n = 588)

Variable	Mean	Std. Dev.	Min	1st Quartile	Median	3rd Quartile	Max
CoE	0.0861	0.0160	0.0355	0.0754	0.0858	0.0959	0.1419
GAP	0.1164	0.0990	0.0000	0.0403	0.0944	0.1699	0.5631
EXP	0.5684	3.3896	0.0000	0.0000	0.0108	0.1005	43.0156
σ_{NI}	0.2056	0.1637	0.0199	0.0970	0.1673	0.2551	1.7593
$\sigma_{CI} - \sigma_{NI}$	0.3003	0.1823	0.0013	0.1702	0.2687	0.3903	1.0667
$\sigma_{FFVI} - \sigma_{CI}$	1.5402	3.4551	0.1341	0.5725	0.8533	1.3545	34.7087

*, **, *** Denotes $p < .10$, $p < .05$, and $p < .01$, respectively, all two-tailed.

CoE denotes the cost of equity capital implied by price per share at the end of year *t*. We estimate CoE using the following model (denoted Equation (13) in the text):

$$CoE = A + \sqrt{A^2 + \frac{eps_1}{P_0} * \left(\frac{eps_2 - eps_1}{eps_1} - (\gamma - 1)\right)}$$

where:

$$A = \frac{1}{2}\left((\gamma - 1) + \frac{dps_1}{P_0}\right)$$

We use consensus I/B/E/S eps forecasts as of the last date for which I/B/E/S compiles forecasts preceding year-end (which they denote as the last statistical date) and CRSP prices as of year-end to estimate Equation (13). As in Gode and Mohanram (2003), we assume γ is equal to $1 + (R_f - 3\%)$.

GAP denotes the absolute value of the excess of interest-bearing assets over interest-bearing liabilities subject to repricing within one year. We obtain GAP from banks' Forms 10-K filed with the SEC and scale it by total assets. If GAP is missing from the Form 10-K we replace it with one-year regulatory GAP obtained from banks' year-end Y9-C regulatory reports.

EXP is total exposure arising from derivatives held at year-end, scaled by total assets. We obtain EXP from Y9-C regulatory reports, and it equals the sum of the notional values of futures contracts, forwards contracts, exchange-written and exchange-purchased options, OTC-written and OTC-purchased options, and swaps.

(*continued on next page*)

TABLE 6 (Continued)

For Columns (a) through (e):

CoE = cost of equity capital estimated as of the end of each year t (t = 2000–2004);

σ_{NI} = firm-specific standard deviation of annual net income expressed as a percentage of average total assets; it is measured over the five-year period ending with each year t (t = 2000–2004);

σ_{CI} = firm-specific standard deviation of annual comprehensive income expressed as a percentage of average total assets; it is measured over the five-year period ending with each year t (t = 2000–2004); and

σ_{FFVI} = firm-specific standard deviation of annual full fair-value income expressed as a percentage of average total assets; it is measured over the five-year period ending with each year t (t = 2000–2004).

Sample sizes vary with each annual cross-section of sample banks with available I/B/E/S data in 2000–2004.

For Column (f):

The variable definitions are the same as in columns (a) through (e), but we estimate the model pooled over years 2000–2004.

For Column (g):

The variable definitions are the same as in columns (a) through (f), except we estimate the volatility metrics over the full nine-year sample period (1996–2004), and we compute CoE as of the end of year 2004.

2003, 2004, and in the five-year pooled estimation. The findings suggest that the incremental volatility in σ_{CI} relative to σ_{NI} is not related to banks' costs of equity capital in any of the models. In Table 6, the most robust result is that the incremental volatility in σ_{FFVI} relative to σ_{CI} is significantly positively related to banks' costs of equity capital in four of the five annual cross-sectional tests (years 2000–2003), in the pooled model, and in the cross-sectional test for 2004 with volatility metrics measured over the full nine-year period. As in our share-price tests, these results are consistent with our prediction that volatility in full-fair-value is a more complete representation of bank risk than volatility in either net income or comprehensive income, and the cost of bank equity capital is increasing in incremental volatility in full-fair-value income.[22]

Robustness Tests

To examine the robustness of our results, we estimated four sets of alternate specifications for the abnormal-earnings-model regressions and obtained results (untabulated) similar to those described in the text and Table 5. First, to verify that the share-price tests we report in Table 5 hold for the subset of bank-years we use for our cost of capital tests in Table 6, we used I/B/E/S analyst forecasts of future earnings as proxies for expected earnings in computing AE. The sample sizes for the cost of capital tests are substantially smaller than those in the share price tests in Table 5 because of the number of banks in the sample that lacked analyst coverage. Despite the reduction in sample size, inferences remain unchanged with this specification, although the R^2s declined in all models.

As a second alternative specification of models (7) through (10), we computed standard deviations of the income metrics by deflating income by book value of equity. We then included these standard deviations of income in a book-value-deflated abnormal-earnings model (i.e., a market-to-book specification of models (7) through (10) in Table 5). Under this specification, inferences, significance levels, and R^2s remain qualitatively similar to those reported in Table 5.

As a third alternative, we used earnings in lieu of abnormal earnings. Inferences remain unchanged and the significance levels and R^2s increased in all models. We did not tabulate these results because we believe that the originally computed abnormal earnings proxy is more theoretically consistent with the residual-income-valuation model. Finally, we further augmented the model by adding an interaction term consisting of book value times each of the income-volatility measures. The coefficients on the book-value/income-volatility-interaction terms are negative and highly significant. The direction and magnitude of the coefficients on the abnormal-earnings/income-volatility interactions are similar to those presented, but are less significant.

V. CONCLUSIONS

Using a sample of 202 commercial banks over 1996–2004, we analyze volatility in three alternate measures of income: reported net income, reported comprehensive income, and a constructed measure of full-fair-value income. We find that the volatility of full-fair-value income is significantly greater than the volatility of net income and comprehensive income for nearly all of our sample banks. Interestingly, we find that average levels of these three income measures over these nine years are similar in magnitude, suggesting that fair-value adjustments in both comprehensive income and full-fair-value income zero out over our sample period.

---

[22] The results in Table 5 indicate that incremental volatility in FFV income negatively moderates the capitalization of abnormal earnings in bank share prices. We attribute these results to volatility in FFV income capturing incremental risk-relevant information but the findings could also be attributable to lower persistence in FFV income. Consistent with our risk explanation, the results we report in Table 4, Panel C provide strong evidence that incremental volatility in FFV income contains incremental risk-relevant information beyond volatility in net income and comprehensive income, and beyond banks' disclosed measures of market-based risks (i.e., repricing gap and derivatives exposure). Further, the results we report in Table 6 indicate that banks' costs of capital are increasing in FFV income volatility, consistent with our risk explanation.

Consistent with bank representatives' criticisms of comprehensive income during the deliberation of SFAS No. 130, we find that reported comprehensive income correlates negatively with unrecognized fair-value changes in financial instruments. In leveling their criticisms, these individuals claimed that comprehensive income would paint an unrealistically harsh picture of banks' core operations and risk-management practices because banks use their total portfolios of financial instruments to manage risk, and other comprehensive income only captures fair-value changes for a small subset of financial instruments (i.e., available-for-sale securities). Contrary to this claim, we find that the partial-fair-value-income accounting and reporting currently mandated in SFAS No. 115 and 130, respectively, substantially *understates* the variability of banks' full-fair-value income. Our results suggest full-fair-value accounting by banks for all financial asset and liability positions reflects more completely the effects of banks' exposure to changes in fair values of financial instruments.

We also investigate the extent to which the volatilities of alternative-income metrics capture market-based risks, including four stock-market-based risk measures (i.e., market-model beta, standard deviation in stock returns, and short-term and long-term interest-rate betas) and two disclosed measures of banks' exposure to market risks (i.e., total derivatives and interest repricing gap). We find volatility in net income is consistently positively correlated with all of those market-risk factors with the exceptions of market-model beta and repricing gap. Volatility in comprehensive income is only consistently positively correlated with standard deviation in returns and short-term interest-rate beta. We find that volatility in FFV income is significantly positively associated with the standard deviation in banks' stock returns and derivatives exposure. We also find that the incremental volatility in FFV income is positively associated with standard deviation in returns, market model beta, and long-term interest-rate beta, beyond the effects of volatility in net income and comprehensive income, and after controlling for repricing gap and derivatives exposure. These results provide support for the conclusion that volatility in FFV income reflects the effects of risk factors that are not completely captured by volatility in net income or comprehensive income, or by banks' disclosures of exposure to market risks.

Across our share-price and cost of capital analyses, we find that FFV-income volatility significantly negatively moderates the capitalization of abnormal earnings in share prices, and is positively associated with implied required rates of return on bank equity capital, consistent with the capital markets pricing FFV-income volatility as risk. We find that volatility in FFV income is negatively associated with abnormal-earnings-capitalization multiples and positively associated with required rates of return even after controlling for volatilities in net income and comprehensive income and after controlling for information in banks' market-risk disclosures regarding repricing gaps and derivatives exposures.

Our findings imply that FFV-income volatility for banks captures important, value-relevant elements of risk that are priced by the capital markets. In addition, FFV-income volatility appears to reflect elements of risk that are not captured by volatility in net income or comprehensive income, or disclosed measures of market risks. We believe these findings are useful to the FASB and the IASB as they consider whether to develop accounting standards to recognize changes in fair values for *all* financial instruments in income. We believe the findings are also useful to capital-markets participants and researchers interested in explaining banks' share prices and risk. Further, we expect these results are useful to bank regulators as they evaluate the role of fair value measurement in bank supervision and regulatory capital monitoring.

We increase the power of our tests by concentrating our analysis on a single industry in which financial instruments are centrally important for operations and risk management. This enhanced statistical-conclusion validity comes at a cost, because our sample of banks may not be representative of the broad cross-section of U.S. and international companies and industries.

APPENDIX

Variable Descriptions and Calculation of Full-Fair-Value Income

This appendix provides descriptions of the accounting variables in this study and sources of data used in their computation.

Accounting Variables Used in the Analysis

- NI = reported net income obtained from SNL DataSource.
- CI = reported comprehensive income obtained from SNL DataSource. CI equals NI plus unrealized fair-value gains and losses on available-for-sale securities plus other comprehensive income items (i.e., foreign currency translation adjustments and pension adjustments).
- FFVI = full-fair-value income. FFVI equals CI plus fair-value adjustments for unrecognized gains and losses on financial instruments (FVA).
- FVA = fair-value adjustments for unrecognized fair-value gains and losses on financial instruments. We compute fair-value adjustments as the change in the excess of fair value over book value of financial instruments. We obtain fair values of recognized financial instruments from SNL DataSource. We obtain fair values of off-balance-sheet items including derivatives from bank holding company regulatory reports (Form Y9-C). We derive fair values of non-term deposits from bank holding company regulatory report data using the approach discussed more completely below. We adjust all fair-value adjustments net of tax effects using the top statutory U.S. Federal corporate income tax rate (35 percent).
- Non-term deposit liabilities = Under current accounting standards, the fair value of non-term deposit liabilities may be assumed to be equal to the face amount plus accrued interest, therefore there are no existing disclosures of the fair values of non-term deposits. Under the proposed standard, the fair value of non-term deposits depends on the estimated timing and amounts of withdrawals, the level of interest rates on other liabilities with similar terms, the cost of servicing the deposits and the financial institution's own credit risk. Valuation techniques estimate the present value of the expected future cash flows by discounting the expected cash flows at a rate equal to that used by the market to price loans of similar risk and effective maturities. When the market rate is higher (lower) than the total cost of deposits, the fair value of the liabilities will be less (greater) than face value.

For purposes of our analysis non-term deposits are valued using a static discounted cash flow approach used by bank regulators together with firm-specific deposit data obtained from bank holding company call reports. The discount equation is shown below:

$$PV_{\textit{Non-term deposits}} = \sum_{t+1}^{360} CF_t df_t \tag{A1}$$

where:

PV = present value of non-term deposit balance;
CF_t = expected cash outflow resulting from that balance in month t; and
df_t = discount factor applicable to month t.

Cash flows are discounted by the zero-coupon treasury rate appropriate to each month plus a spread. The spread is a parameter that calibrates demand deposit prices to observed prices for deposit purchase/assumption transactions and is obtained from the U.S. Treasury Department website (http://www.treasury.ots.gov).

Cash flows in the discount equation are given by the following:

$$CF_t = NIC(B_{t-1}) + [(1 + r_t)B_{t-1} - B_t] \tag{A2}$$

where:

CF_t = cash flow in month t;
NIC = rate of noninterest cost of retail deposits;
B_t = demand deposit balance at end of month t; and
r_t = deposit rate in month t.

The deposit rate (r_t) in any month is assumed to be a function of three factors: (1) the rate paid by the institution in the prior month, (2) projected market rates for the current month, and (3) the difference between the rates paid and market rates, which are assumed to gradually converge in equilibrium. Specifically:

$$r_t = r_{t-1} + c(r_{t-1} - r_{t-2}) + d(R_t - R_{t-1}) + e(R_{t-2} - R_{t-3}) + C_{t-1} \tag{A3}$$

where:

r_t = deposit rate in month t;
R_t = implied forward short-term rates in month t;
C_t = adjustment toward equilibrium rate = $f(r_{t-1} - E_{t-1})$, if $r_{t-1} \geq E_{t-1}$
$g(r_{t-1} - E_{t-1})$, if $r_{t-1} < E_{t-1}$; and
E_t = equilibrium interest rate for demand deposits = $a + bRt$.

The parameters (a–g) in the above equations are determined using historical data. For example, a and b are parameters obtained from regressing deposit rates on one-month LIBOR while c, d, and e are obtained from a regression of current rates on lagged changes in rates. Estimates for all parameters are obtained from the U.S. Treasury Department website (http://www.treasury.ots.gov).

Finally, balances in non-term deposit accounts are assumed to decline over time with the rate of decline influenced by the relation of the rate paid on the account to market rates (the ratio of r_t/R_t). The decline is nonlinear and assumed to take the following form:

$$B_t = B_{t-1}(h + iarctan(k + jr_t/R_t) + lr_t)^{1/12} \tag{A4}$$

where:

B_t = non-term deposit balance at end of month t;
$arctan$ = arctangent function;
r_t = interest rate offered on deposits in month t by institution i; and
R_t = implied forward short-term rates in month t (annual percentage form).

As with the other equations, all parameters in Equation (A4) (h–l) are based on aggregate historical balance migration data obtained from the U.S. Treasury Department. The short-term market rate series used is one-month LIBOR. Annual deposit servicing cost parameters were obtained from the Office of Thrift Supervision "Selected Asset and Liability Price Tables" (http://www.ots.treas.gov). Because firm-specific transaction account interest costs were not available, the interest component of cost was estimated by matching the quintile rank of the financial institution based on total deposit cost to the mean interest rate for that quintile of transaction deposits from the Selected Asset and Liability Price Tables.

REFERENCES

Ahmed, A., and C. Takeda. 1995. Stock market valuation of gains and losses on commercial banks' investment securities: An empirical analysis. *Journal of Accounting & Economics* 20 (2): 207–236.

Baginski, S., and J. Wahlen. 2003. Residual income risk, intrinsic values and share prices. *The Accounting Review* 78 (January): 327–351.

Barth, M. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69 (January): 1–25.

———, W. Landsman, and J. Wahlen. 1995. Fair value accounting: Effects on banks' earnings volatility, regulatory capital, and value of contractual cash flows. *Journal of Banking and Finance* 19: 577–605.

———, W. Beaver, and W. Landsman. 1996. Value-relevance of banks' fair value disclosures under SFAS No. 107. *The Accounting Review* (October): 513–537.

———, ———, and ———. 2001. The relevance of the value relevance literature for financial accounting standard setting: Another view. *Journal of Accounting and Economics* (31): 77–104.

Beaver, W., P. Kettler, and M. Scholes. 1970. The association between market-determined and accounting-determined measures of risk. *The Accounting Review* (October): 654–682.

Botosan, C., and M. Plumlee. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40 (March): 21–40.

———, and ———. 2005. Assessing alternative proxies for the expected risk premium. *The Accounting Review* (January): 21–53.

Cahan, S. F., S. M. Courtenay, P. L. Gronewoller, and D. Upton. 2000. Value-relevance of mandated comprehensive income disclosures. *Journal of Business, Finance and Accounting* 27 (November/December): 1273–1301.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56 (October): 1629–1666.

Dean, T. J. 2000. Banks need to write on fair value. *ABA Bankers News* (April 18): 3.

Dhaliwal, D., K. R. Subramanyam, and R. Trezevant. 1998. Is comprehensive income superior to net income as a measure of firm performance? *Journal of Accounting and Economics* 26 (January): 43–67.

Easton, P., G. Taylor, P. Shroff, and T. Sougiannis. 2002. Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research* 40 (3): 657–677.

Eccher, A., K. Ramesh, and R. Thiagarajan. 1996. Fair value disclosures by bank holding companies. *Journal of Accounting & Economics* 22: 79–117.

Elgers, P. 1980. Accounting-based risk predictions: A re-examination. *The Accounting Review* 55 (3): 389–408.

Fama, E., and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47: 427–465.

Feltham, G. A., and J. Ohlson. 1999. Residual earnings valuation with risk and stochastic interest rates. *The Accounting Review* 74 (2): 165–183.

Financial Accounting Standards Board (FASB). 1991. *Disclosures about Fair Value of Financial Instruments*. Statement of Financial Accounting Standards No. 107. Norwalk, CT: FASB.

———. 1993. *Accounting for Certain Investments in Debt and Equity Securities*. Statement of Financial Accounting Standards No. 115. Norwalk, CT: FASB.

———. 1997. *Reporting Comprehensive Income*. Statement of Financial Accounting Standards No. 130. Norwalk, CT: FASB.

———. 1998. *Accounting for Derivative Instruments and Hedging Activities*. Statement of Financial Accounting Standards No. 133. Norwalk, CT: FASB.

Flannery, M. and C. James. 1984. The effect of interest rate changes on the common stock returns of financial institutions. *Journal of Finance* (39): 1141–1153.

Gebhardt, W., C. M. C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39 (June): 135–176.

Gode, D., and J. Ohlson. 2004. Accounting-based valuation with changing interest rates. *Review of Accounting Studies* 9 (4): 419–449.

Hirst, E., P. Hopkins, and A. Yen. 2002. A content analysis of the comprehensive income exposure draft comment letters. Working paper, The University of Texas at Austin.

———, ———, and J. Wahlen. 2004. Fair values, income measurement, and bank analysts' risk and valuation judgments. *The Accounting Review* 79 (2): 453–472.

Hodder, L. 2005. Relevance of market risk disclosures by commercial banks. Working paper, Indiana University.

Holthausen, R., and R. Watts. 2001. The relevance of the value relevance literature for financial accounting standard setting. *Journal of Accounting and Economics* (31): 3–75.

Joint Working Group of Banking Associations on Financial Instruments. 1999. *Accounting for Financial Instruments for Banks*. New York, NY: Self published.

Joint Working Group of Standard Setters. 2000. *Recommendations on Accounting for Financial Instruments and Similar Items*. Norwalk, CT: FASB.

Nelson, K. 1996. Fair value accounting for commercial banks: An empirical analysis of SFAS No. 107. *The Accounting Review* 71 (April): 161–182.

Ohlson, J. A. 1995. Earnings, book values and dividends in equity valuation. *Contemporary Accounting Research* (Spring): 661–687.

———, and B. E. Nauroth-Juettner. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10 (June–September): 349–366.

Petroni, K., and J. Wahlen. 1995. Fair values of equity and debt securities and share prices of property-liability insurers. *Journal of Risk and Insurance* 62 (December): 719–738.

Riffe, S., 1996. The valuation of off-balance-sheet financial instrument disclosures in the banking industry. In *Derivatives, Regulation and Banking: Advances in Finance, Investment and Banking Series*. Amsterdam, The Netherlands: North-Holland.

Ryan, S. 1997. A survey of research relating accounting numbers to systematic equity risk, with implications for risk disclosure policy and future research. *Accounting Horizons* (June): 82–95.

Schrand, C. 1997. The asscciation between stock-price interest rate sensitivity and disclosures about derivative instruments. *The Accounting Review* (72): 87–109.

Venkatachalam, M. 1996. Value-relevance of banks' derivative disclosures. *Journal of Accounting and Economics* (27): 327–355.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 377–397

Using Counter-Explanation to Limit Analysts' Forecast Optimism

Kathryn Kadous
Emory University

Susan D. Krische
University of Illinois at Urbana–Champaign

Lisa M. Sedor
University of Notre Dame

ABSTRACT: Prior research demonstrates that forecast optimism is, in part, a consequence of analysts' cognitive reactions to the scenarios managers use to communicate future plans. In two experiments, we examine whether counter-explanation (explaining why managers' plans could fail) reduces scenario-induced optimism. We find that when compared to analysts not asked to generate counter-explanations, analysts who complete the relatively easy task of generating few counter-explanations make less optimistic forecasts, but analysts who complete the relatively difficult task of generating many counter-explanations do not. Results demonstrate the usefulness of a cognitive, theory-based mechanism for reducing forecast optimism and suggest a boundary condition for the use of that mechanism.

Keywords: *counter-explanation; availability; EPS; forecast optimism.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

In a move applauded by Wall Street analysts, Coca-Cola Co. announced that it would no longer provide quarterly earnings guidance because such short-run guidance "prevents a more meaningful focus on the strategic initiatives that a company is taking to

---

We thank Michael Bamber, Walt Blacconiere, Bryan Church, Peter Easton, Paquita Davis-Friday, Craig Emby, Ken Gaver, Jackie Hammersley, Gary Hecht, Lisa Koonce, Marlys Lipe (editor), Laureen Maines, Molly Mercer, Jeff Miller, Fred Mittelstaedt, Mike Morris, Jon Perkins, Auke Plantinga, Jamie Pratt, Dave Ricchiute, Peter Roberts, Jerry Salamon, Tom Schaefer, Margaret Shackell-Dowell, Jim Seida, Nathan Stuart, Mike Tiller, Sandra Vera-Muñoz, Michael Williamson, Alex Yen, two anonymous reviewers, workshop participants at the 9th Behavioral Decision Making Research in Management Conference, the 2004 American Accounting Association Annual Meeting, the 2004 Canadian Academic Accounting Association annual meeting, DePaul University's People & Money—The Human Factor in Financial Decision-Making Conference, Indiana University, University of Florida, University of Georgia, University of Montana, and University of Notre Dame for helpful comments; David Cantos, Jim Frederickson, Oliver Li, Jeff Miller, Dave Ricchiute, Peter Roberts, and Sandra Vera-Muñoz for help with our materials; and Andrew Litten, Aakash Ravanat, David Reppenhagen, Jane Thayer, and Charlie Yuan for research assistance.

Editor's note: This paper was accepted by Marlys Gascho Lipe, Editor.

Submitted October 2004
Accepted July 2005

build its business and to succeed over the long run" (*Wall Street Journal Online* 2002). In connection with this shift away from providing quarterly earnings guidance, investor-relations and other executives are moving toward guiding investors' expectations for annual earnings (National Investor Relations Institute [NIRI] 2003; Zuckerman 2005). Managers eliminating short-run earnings guidance to promote a focus on long-run performance will likely increase their use of qualitative or narrative communications regarding their plans and future earnings expectations. Yet, managers' use of detailed, causal narratives (i.e., scenarios) to communicate plans to improve future earnings facilitates analysts' ability to envision how and why these plans will improve future firm performance, resulting in optimistic analysts' earnings forecasts (Sedor 2002). In this paper, we test whether generating counter-explanations (i.e., reasons why managers' plans could fail) mitigates optimism in analysts' forecasts. We also provide evidence regarding a boundary condition for the use of counter-explanation in this domain.

It is likely that optimism in analysts' forecasts is costly to investors because investors do not appear to unravel analysts' forecast optimism completely (e.g., Dechow et al. 2000; Bradshaw et al. 2003). In fact, investors who use analysts' forecasts to estimate firm value could improve the quality of their valuation estimates by correcting for analysts' optimism (Frankel and Lee 1998). Additionally, analysts appear to use their annual earnings forecasts to develop their recommendations (e.g., Bradshaw 2004), and the accuracy of analysts' forecasts is associated with the profitability of their contemporaneous recommendations (Loh and Mian 2005). Consequently, reducing optimism in analysts' earnings forecasts also could result in higher returns for investors relying on analysts' recommendations.

Drawing on theory in psychology, we propose that asking analysts to generate counter-explanations will reduce forecast optimism. Our theory further suggests that the ease or difficulty that analysts experience while generating counter-explanations will influence the effect of the counter-explanation mechanism on analysts' forecasts. Consistent with availability, we expect that the ease or difficulty of generating counter-explanations will act as a heuristic cue and qualify the influence of counter-explanation content on analysts' forecasts.[1] Specifically, we hypothesize that optimism will be reduced when analysts generate counter-explanations with relative ease, but not when analysts generate counter-explanations with relative difficulty.

In Experiment 1, we provide professional analysts with summarized financial information for a firm and a scenario describing managers' plans to improve future earnings. From this information, analysts in the control condition provide one-year- and two-year-ahead annual earnings forecasts. Analysts in the experimental conditions list reasons why managers' plans could fail (i.e., generate counter-explanations) prior to providing their forecasts. We manipulate the availability of counter-explanations by asking analysts to generate either few (which is relatively easy) or many (which is relatively difficult) counter-explanations. We find that analysts who perform the relatively easy task of generating few counter-explanations make annual earnings forecasts that are lower than the forecasts made by analysts in the control condition. In contrast, analysts who perform the relatively difficult task of generating many counter-explanations make annual earnings forecasts that do not differ from the forecasts made by analysts in the control condition. We find this result even though analysts in the generate many condition generate more reasons why managers' plans could fail than do analysts in the generate few condition.

---

[1] Individuals employ the *availability heuristic* when they estimate probabilities or frequencies by the ease with which relevant information can be brought to mind (Tversky and Kahneman 1973). The term *availability* reflects the use of two types of mental operations that bring information to mind: direct recall and ease of generation (Hoch 1984). We focus our investigation on availability as the ease (or difficulty) of generation.

In Experiment 2, we provide additional evidence that the availability of counter-explanations, and not differences in counter-explanation content, explains Experiment 1 results. In particular, in Experiment 2, we replicate Experiment 1 with M.B.A. student participants and add two experimental conditions in which all counter-explanations (regardless of number) are equally available to participants. Each participant in these new conditions reads a reproduction of the counter-explanations generated by another participant instead of generating his or her own. Specifically, each participant in the read few (read many) condition is presented with the set of counter-explanations created by a matched participant in the generate few (generate many) condition. This yoked design ensures that the quality and order of counter-explanations are identical between generate and read conditions.

Consistent with Experiment 1 results, we find in Experiment 2 that participants asked to perform the relatively easy task of generating few counter-explanations make annual earnings forecasts that are lower than the forecasts made by participants in the control condition, but that participants asked to perform the relatively difficult task of generating many counter-explanations make annual earnings forecasts that do not differ from the forecasts made by participants in the control condition. We also find that participants who read the counter-explanations generated by other participants, regardless of the number of counter-explanations read, make annual earnings forecasts that are lower than the forecasts made by participants in the control condition. If differences in counter-explanation content explained results, then forecasts made by read condition participants would mirror forecasts made by generate condition participants (Wanke et al. 1996). In contrast, forecasts made by participants who read *either* the few *or* many counter-explanations generated by matched participants are lower than the forecasts made by participants who neither generated nor read counter-explanations, supporting our theory that the availability of counter-explanations explains reported results.

Our research makes three contributions. First, we identify and test counter-explanation as a mechanism to reduce optimism in analysts' forecasts. Prior research suggests that optimism is due, in part, to managers' use of scenarios to communicate their plans to improve future firm performance (Sedor 2002). Because managers are likely to continue using scenarios to communicate their plans, recipients of those communications bear the burden of taking measures to control their cognitive reactions to the scenarios. The counter-explanation mechanism is practical and could be implemented by analysts and others who wish to avoid optimism, but who rely on causal narratives that describe positive future outcomes when developing forecasts.

Second, our results provide insight into potential causes of forecast optimism among analysts. In particular, results are consistent with analysts failing to consider spontaneously a small number of counter-explanations when contemplating scenarios describing managers' plans. To the extent that scenario-induced optimism is manifest in analysts' forecasts of annual earnings, failing to consider counter-explanations may be one factor contributing to the optimism. Finally, we demonstrate an important boundary condition for the use of counter-explanation as a debiasing mechanism. In both experiments, we find that the ability to generate counter-explanations with ease is critical to the effectiveness of this mechanism.

We use an experimental methodology for our investigation because archival data neither provide information regarding the cognitive forecasting processes used by professional analysts nor a control condition against which to compare the effects of manipulating counter-explanation availability among analysts. Additionally, we hold constant investment banking relationships, which enables us to examine how forecast optimism is affected by counter-explanation in the absence of economic incentives. A limitation of our experimental approach is that we cannot measure absolute levels of forecast optimism, and thus the evidence is limited to *relative* levels of optimism across experimental conditions.

The remainder of this paper is organized as follows. Section II analyzes relevant literature and develops the first set of hypotheses. Section III describes the design, method, and results for Experiment 1. Section IV develops the second set of hypotheses and describes the design, method, and results for Experiment 2. The final section summarizes and concludes.

II. LITERATURE ANALYSIS AND HYPOTHESES

Forecast Optimism and Scenarios

Optimism in annual earnings forecasts is statistically and economically significant (e.g., Darrough and Russell 2002; Duru and Reeb 2002) and increases over the forecast horizon: analysts' two-year-ahead forecasts are relatively more optimistic than their one-year-ahead forecasts (DeBondt and Thaler 1990; Frankel and Lee 1998; Bradshaw et al. 2003). Although optimism in quarterly earnings forecasts decreases as the earnings announcement date approaches (e.g., Brown 1997, 2001; Matsumoto 2002; Richardson et al. 2004), analysts' forecasts of annual earnings are often used to make investment decisions and to estimate firm value.[2]

Much, but not all, of the existing literature has focused on incentive-based explanations for analysts' forecast optimism.[3] Recent research, however, suggests that analysts are unlikely to issue intentionally optimistic forecasts (Eames and Glover 2003; Irvine 2004).[4] Consistent with forecast optimism being driven by something other than analysts' responses to economic incentives, market analysts (Darrough and Russell 2002), analysts employed by independent research firms (Jacob et al. 2003), and students (Affleck-Graves et al. 1990) also make optimistic earnings forecasts.

We focus our investigation on a key aspect of analysts' information environment: the detailed, causal narratives (i.e., scenarios) that managers commonly use to communicate their plans to improve future firm performance (via conference calls, press releases, and other communications). Optimism in analysts' annual earnings forecasts derives in part from analysts' cognitive reactions to these scenarios (Sedor 2002). When analysts contemplate managers' plans in scenarios, they envision in concrete detail how managers will implement plans and, ultimately, how and why those plans will improve future firm performance (Sedor

---

[2] For example, empirical applications of the residual income model often use analysts' forecasts as proxies for the market's expectations of future earnings (e.g., Frankel and Lee 1998; Dechow et al. 1999; Lee et al. 1999; Liu and Thomas 2000; Sougiannis and Yaekura 2001; Ali et al. 2003). Bradshaw (2004) suggests that analysts' recommendations reflect a simple heuristic based on forecasts of annual earnings and long-term earnings growth. Ohlson and Juettner-Nauroth (2005) also model firm value as a function of annual earnings and earnings growth forecasts.

[3] Analysts may provide optimistic forecasts in order to maintain or gain access to information (Francis and Philbrick 1993; Krishnan and Sivaramakrishnan 1998; Lim 2001), to motivate trades (Kim and Lustgarten 1998), and to obtain investment banking business (Dugar and Nathan 1995; Hunton and McEwen 1997; Lin and McNichols 1998; Dechow et al. 2000; Bradshaw et al. 2003). Other explanations include justifying favorable recommendations (Eames et al. 2002), failing to adjust for earnings management (Abarbanell and Lehavy 2003), and choosing to report only on stocks that are viewed favorably (McNichols and O'Brien 1997).

[4] This is consistent with research suggesting that analysts' reputations depend on accurate forecasts and recommendations (Michaely and Womack 1999; Loh and Mian 2005). Additionally, greater forecast accuracy decreases the likelihood of termination (Mikhail et al. 1999) and increases the likelihood of moving from a lower to a higher status brokerage firm (Hong and Kubik 2003).

2002). Envisioning ways in which managers' plans will improve future firm performance increases the availability of those plans and the associated positive outcomes to analysts (i.e., the plans and resultant performance improvements are easy to bring to mind). Individuals use the ease of bringing information to mind as a heuristic cue when making judgments about probabilities or frequencies (Tversky and Kahneman 1973; Kahneman and Tversky 1982; Kahneman and Lovallo 1993). Consequently, analysts will infer from the availability of managers' plans and the associated positive outcomes that future performance improvements are likely, and optimism in analysts' forecasts will result.

Scenarios and Counter-Explanation

Effective debiasing of scenario-induced forecast optimism requires developing a mechanism that addresses the optimism's cognitive source (Arkes 1991; Kennedy 1993, 1995; Bonner et al. 1996). For instance, the mechanism could increase the availability of causal chains of events leading to alternative (less positive) outcomes. One such mechanism is counter-explanation.

Prior research conducted in the context of analytical review tasks demonstrates that counter-explanation reduces auditors' likelihood assessments that unusual account balance fluctuations were driven by hypothesized causes. Heiman (1990) showed that both generating and reading counter-explanations reduced auditors' likelihood assessments for an experimenter-suggested cause. Koonce (1992) had a group of auditors generate counter-explanations to a manager-suggested cause for an unusual account balance fluctuation. She found that the auditors who generated counter-explanations rated the manager's explanation as less likely than did auditors who did not generate counter-explanations. Similarly, we expect that generating counter-explanations to managers' plans to improve future firm performance (that is, generating reasons why managers' plans could fail) will reduce analysts' beliefs that the managers' plans are likely to be successful and will mitigate analysts' forecast optimism.[5]

Heiman (1990) found that auditors who considered five counter-explanations made lower likelihood assessments than did auditors who considered two counter-explanations. This finding suggests that the larger the number of counter-explanations considered by analysts, the greater the anticipated effect of the counter-explanations on forecasts. However, we draw on more recent psychological theories to develop hypotheses regarding a boundary condition for the use of counter-explanation to mitigate analysts' forecast optimism.

Boundary Condition: The Experienced Ease or Difficulty of Generating Counter-Explanations

We apply the theoretical framework of availability to understand the expected boundary condition for the use of counter-explanation as a debiasing task. There are two sources of information associated with items, such as counter-explanations, that are relevant to judgments and decisions: the information content of the items and the experienced ease or difficulty with which those items are brought to mind (i.e., availability) (Hoch 1984; Schwarz et al. 1991; Schwarz 1998). The availability of information often affects the influence of information content on judgment (Sanna et al. 2002; Wanke et al. 1997; Schwarz et al. 1991). This availability effect is most likely when individuals consider themselves to be knowledgeable in a domain because they then believe that the ease or difficulty they experience when bringing information to mind is informative (Schwarz 1998).

---

[5] To the extent that analysts' professional responsibilities include skepticism, such skepticism could encourage spontaneous counter-explanation and would bias against our ability to detect an effect when analysts are asked to generate counter-explanations (Heiman 1990; Koonce 1992).

Consistent with availability affecting the influence of information content on judgment, Sanna et al. (2002) found that asking participants to generate two thoughts about how a war could have turned out differently (which participants found easy to do) attenuated hindsight (i.e., reduced their exaggerated assessments of the inevitability of past events). However, asking participants to generate ten thoughts (which participants found difficult to do) exacerbated hindsight effects, despite the fact that the latter group identified significantly more thoughts about how the war's outcome could have occurred differently than did the former group.

Similarly, we expect that the availability of counter-explanations (i.e., the ease or difficulty of their generation) will influence whether counter-explanation mitigates scenario-induced optimism. If analysts generate counter-explanations with relative ease, then we expect that analysts will use the availability of the counter-explanations as a heuristic cue suggesting that future performance improvements are unlikely (because it is easy to bring to mind ways that managers' plans could fail to improve future performance). As a result, forecast optimism will be reduced (Hypothesis 1). On the other hand, if analysts generate counter-explanations with relative difficulty, then we expect that analysts will use the lack of available counter-explanations as a heuristic cue suggesting that future performance improvements are likely (because it is difficult to bring to mind ways that managers' plans could fail to improve future performance). In this case, forecast optimism will not be reduced (Hypothesis 2) and may even be increased.[6] However, with additional counter-explanation content, it is difficult to predict the net effect of requiring the generation of many counter-explanations. Thus, in recognition of the competing influences of counter-explanation availability and additional counter-explanation content, our second hypothesis is not directional. We test the following predictions:

> **H1:** Analysts who are asked to *generate few* counter-explanations of managers' plans to improve future earnings will make earnings forecasts that are *less* optimistic than the earnings forecasts made by analysts who are not asked to generate counter-explanations.
>
> **H2:** Analysts who are asked to *generate many* counter-explanations of managers' plans to improve future earnings will make earnings forecasts that *do not differ* from the earnings forecasts made by analysts who are not asked to generate counter-explanations.

II. EXPERIMENT 1

In Experiment 1, we examine whether generating counter-explanations can reduce scenario-induced optimism in analysts' earnings forecasts and whether the availability of counter-explanations influences the effectiveness of this mechanism. We use a 1 × 3 between-subjects design with the number of counter-explanations participants are asked to generate (none, few, many) as the independent variable. For presentation purposes, we calculate our dependent variable by subtracting current earnings per share (EPS) as provided to participants from participants' annual EPS forecasts for one year and two years ahead.

---

[6] It is unclear, *ex ante*, whether the effects of counter-explanation observed by Sanna et al. (2002) in the context of debiasing hindsight (exaggerated *retrospective* likelihood assessments) and Heiman (1990) and Koonce (1992) in the context of reducing reliance on initial hypotheses in diagnosis (also exaggerated *retrospective* likelihood assessments) will be observed in the context of debiasing forecast optimism. Thus, we test whether counter-explanation tasks will influence *prospective* judgments in a way that mirrors their effect on *retrospective* judgments.

Design and Method

Participants

Participants are 59 financial analysts recruited via telephone calls placed by the authors using *Nelson's Directory of Investment Research* (2003). We awarded a prize of $500 to one participant via random drawing as a token of our appreciation. Participants reported a mean (median) of 11 (8) years of experience. Ninety-three percent of participants identified themselves as sell-side analysts, 36 percent as certified financial analysts (CFAs), and 37 percent as being named by *Institutional Investor* as All-Star Analysts. Participants reported a mean (median) of 1.4 (1.0) industry specializations, with the most common being entertainment (19 percent), telecommunications (19 percent), health (10 percent), and transportation (10 percent).[7] We used a predetermined randomization plan to assign participants to experimental conditions as they were recruited. We provided each participant with an electronic link to the experimental materials for his or her assigned condition, and participants completed the study online.[8]

Case Materials and Procedures

Case materials were adapted from the loss-firm/scenario-structure condition used in Sedor (2002).[9] We focus on a loss firm for several reasons. First, the effect of scenario-induced optimism on analysts' earnings forecasts is larger for firms reporting losses than it is for firms reporting profits (Sedor 2002). Correspondingly, the most optimistic annual earnings forecasts issued by analysts are associated with firms that report losses in the prior year (Klein 1990; Ali et. al. 1992; Dowen 1996) or that have weak prior earnings performance (Abarbanell and Bernard 1992; Elgers and Lo 1994). Finally, firms reporting consecutive losses are likely to discontinue earnings guidance (Graham et al. 2005). Thus, we believe that managers of loss firms are more likely to use qualitative or narrative communications when presenting their plans and future expectations.

To avoid the influence of economic incentives on forecasts, we instructed each participant to assume that he or she was initiating coverage of a company (Villa Fiore Winery) that did not have (and did not contemplate) an investment banking relationship with the participant's firm. Participants began by reading three years of summarized financial information (including net losses and negative EPS for the current and prior two years), a brief description of the company, and product information. Participants then provided initial forecasts of annual EPS for one year and two years ahead.[10] Participants had access to the summarized financial information while making their initial forecasts. Next, participants read an excerpt from a conference call held after the company's annual earnings announcement. The conference-call excerpt included the CEO's discussion of the industry, the company, and a description of managers' plans to improve future earnings in scenario structure. Participants then generated counter-explanations (if applicable) and made final EPS forecasts. Electronic links provided access to the summarized financial information and to the conference-call excerpt (but not the counter-explanations, if applicable) while participants were making their final forecasts. Finally, participants completed a post-experimental questionnaire.

---

[7] Including either the author who recruited the analyst, years of experience, CFA designation, or All-Star Analyst rating in the models when testing our hypotheses does not change reported inferences. Additionally, inferences remain unchanged if data from buy-side analysts or the analyst who indicated a beverage specialization are excluded.

[8] Although approximately equal numbers of participants were assigned to each condition, and approximately equal numbers of participants in each condition accessed the study, participants in the generate many condition were somewhat less likely to complete both sets of forecasts than were participants in the other two conditions ($\chi^2 = 5.12$, $p = 0.08$), and thus the cell size for the generate many condition is the smallest. However, of the 16 analysts who dropped out before completing the study, 12 did so before making initial forecasts. These participants could not have seen the counter-explanation manipulation because it took place after initial forecasts were elicited. Fisher's Exact Test also reveals that the timing of dropouts does not differ by experimental condition ($p = 0.73$). Thus, it is unlikely that the manipulation caused differences in drop out rates across experimental conditions.

[9] Sedor's (2002) materials were constructed from Forms 10-K filed by companies in the wine industry. Advantages of these materials include (1) at (and before) the time the experiment was conducted, the press had not reported any noteworthy news events for the wine industry (e.g., excessive supplies, recalls); (2) the wine industry is not followed by a large number of analysts (minimizing any potential advantage of industry-specific expertise); and (3) these materials induced scenario thinking among financial analysts in Sedor (2002). We updated all materials for recent industry events, but maintained their original structure.

Independent Variable

We manipulated the availability of counter-explanations (i.e., the ease or difficulty of their generation) by asking participants to generate varying numbers of counter-explanations after they read the conference-call excerpt, but before they made final forecasts. Specifically, we provided participants in the generate few (generate many) condition with the following instruction:

> Imagine that firm policy requires you to document (as part of your analysis) reasons why managers' plans COULD FAIL. Below, please provide TWO (TWELVE) reasons why the plans presented by Villa Fiore's managers during the conference call COULD FAIL to increase sales and profitability as expected.

The online instrument contained either two or 12 blank lines for participants' counter-explanations, depending on experimental condition. The instrument was programmed such that participants were required to input the requested number of counter-explanations before continuing. Control condition participants were not explicitly asked to generate counter-explanations; instead they made final EPS forecasts immediately after reading the conference-call excerpt.

Dependent Measures

Participants provided initial and final forecasts of annual EPS for one year and two years ahead. For presentation purposes, we subtract current (year 0) EPS of −$0.64 from participants' EPS forecasts to arrive at participants' forecast increases in EPS. The dependent measures are participants' final forecast increases in EPS for both one year (year 1) and two years ahead (year 2). We use participants' initial forecast increases in EPS as covariates in the analyses. Table 1, Panel A reports descriptive statistics for Experiment 1, including means and least squares (adjusted) means for final forecast increases in EPS for years 1 and 2.

---

[10] Participants also forecasted five-year long-term growth rates (LTG). The pattern of LTG matches that of EPS forecasts, but the manipulations do not explain significant variation in LTG (smallest p = 0.08). We believe this is due to the inherent difficulty in forecasting long-term growth rates for firms with negative earnings and thus, LTG is not discussed further.

TABLE 1
Experiment 1: Professional Analysts' EPS Forecasts

Panel A: Descriptive Statistics: Mean [standard deviation]

		Year 1			Year 2		
Condition	**n**	**Initial Forecast Increase in EPS**	**Final Forecast Increase in EPS**	**ANCOVA Adjusted Final Forecast Increase in EPS[a]**	**Initial Forecast Increase in EPS**	**Final Forecast Increase in EPS**	**ANCOVA Adjusted Final Forecast Increase in EPS[a]**
Control	17	0.23 [0.18]	0.30 [0.22]	0.29	0.57 [0.35]	0.67 [0.31]	0.62
Generate Few	24	0.21 [0.15]	0.16 [0.20]	0.17	0.45 [0.24]	0.44 [0.31]	0.46
Generate Many	15	0.23 [0.18]	0.25 [0.24]	0.25	0.46 [0.27]	0.52 [0.29]	0.54

Panel B: Planned Contrast Tests of Hypotheses

Hypothesis Contrast	**Year 1 t-statistic**	**Year 1 p-value**	**Year 2 t-statistic**	**Year 2 p-value**
H1: Lower relative optimism for generate few versus control	2.09	0.02	1.92	0.03
H2: No difference in relative optimism for generate many versus control	0.68	0.50	0.95	0.35

---

[a] The final forecast increase in EPS is adjusted statistically for the initial forecast increase in EPS by including this as a covariate in the ANCOVA model. The adjusted means are reported here.

Panel A provides descriptive statistics by experimental condition. "Year *t* final forecast increase in EPS" is our dependent variable and is defined as the final (post-manipulation) forecast for year *t* ($t = 1$ or 2) less the year 0 EPS of −$0.64 as provided to participants. In our analyses, we adjust for the influence of participants' initial (pre-manipulation) forecasts using analysis of covariance (ANCOVA) and thus, we report the ANCOVA least squares adjusted means in Panel A. Participants made initial EPS forecasts after reading summarized financial information and made final EPS forecasts after reading a conference-call excerpt (which presented information about managers' plans to improve future earnings in scenario structure) and generating counter-explanations (if applicable). Participants in the generate few (generate many) condition were asked to list 2 (12) reasons why managers' plans could fail before making final forecasts. Participants in the control condition were not asked to generate counter-explanations before making final forecasts.

Panel B provides tests of hypotheses comparing the forecast increases in EPS of participants in experimental conditions to those of participants not asked to generate counter-explanations (i.e., participants in the control condition) to examine whether the availability of counter-explanations (i.e., the ease or difficulty experienced during the counter-explanation task) affects the relative amount of scenario-induced optimism. If the forecast increases in EPS in experimental conditions are less than the forecast increases in EPS in the control condition, then experimental condition participants are *relatively less optimistic*. P-values are one-tailed for the directional prediction in H1 and two-tailed for H2.

Results

Preliminary Analyses and Manipulation Check

We exclude three participants' forecast increases in EPS from the analyses because they differed from sample means by more than four standard deviations, and diagnostic procedures indicated that they are influential outliers. Reported analyses use the remaining 56 observations.

We use forecasts made by control condition participants as our baseline for measuring changes in scenario-induced optimism. We expect forecasts made by participants who contemplate managers' plans to improve future firm performance in scenarios to exhibit optimism (as in Sedor [2002]). Consistent with this, control condition participants gave final forecast increases in EPS after reading the conference-call excerpt that exceeded their initial forecast increases in EPS by \$0.07 for year 1 ($t_{16} = 1.39$, $p = 0.09$) and by \$0.10 ($t_{16} = 2.01$, $p = 0.03$) for year 2.[11]

In a post-experimental question, we asked participants who generated counter-explanations to rate how difficult it was for them to do so using a scale ranging from 1 (not at all difficult) to 9 (extremely difficult). Consistent with a successful manipulation, the difficulty ratings provided by participants asked to generate 12 counter-explanations were significantly greater than the difficulty ratings provided by participants asked to generate two counter-explanations (means [standard deviations] = 3.67 [2.16] and 1.88 [1.30] for the generate many and generate few conditions, respectively; $t_{37} = 3.25$, $p < 0.01$).[12]

Tests of Hypotheses

Table 1, Panel B reports results of the planned contrasts used to test our hypotheses. Participants asked to generate few counter-explanations gave significantly less optimistic forecast increases in EPS than did control condition participants ($t_{55} = 2.09$, $p = 0.02$ for year 1 and $t_{55} = 1.92$, $p = 0.03$ for year 2), supporting H1. Participants asked to generate many counter-explanations gave forecast increases in EPS that did not differ from those of control condition participants ($t_{55} = 0.68$, $p = 0.50$ for year 1 and $t_{55} = 0.95$, $p = 0.35$ for year 2), supporting H2.[13]

Theory predicts that the availability of counter-explanations (i.e., the ease or difficulty of their generation) will influence the effectiveness of the counter-explanation task, rather than the number of counter-explanations generated per se. Thus, we repeat the main analyses using a median split of the difficulty of generation measure (manipulation check) as the independent variable.[14] Participants reporting that it was relatively easy to generate counter-explanations gave marginally less optimistic forecast increases in EPS than did control condition participants for year 1 (adjusted means = 0.29 for control and 0.20 for lower difficulty; $t_{52} = 1.45$, $p = 0.08$) and significantly less optimistic forecast increases in EPS for year 2 (adjusted means = 0.62 for control and 0.41 for lower difficulty; $t_{52} = 2.67$, $p < 0.01$), providing partial support for H1. Participants reporting higher difficulty on the counter-explanation task gave forecast increases in EPS that did not differ from those of control condition participants for either year (adjusted means = 0.29 and 0.19 for control and higher difficulty, respectively, t_{52} = 1.54, p = 0.13 for year 1; and adjusted means = 0.62 and 0.63 for control and higher difficulty, respectively, $t_{52} < 0.01$, p = 0.94 for year 2), supporting H2.

---

[11] Control condition participants' unadjusted mean final forecast increases in EPS for years 1 and 2 are \$0.30 and \$0.67, respectively. Despite the changes in analysts' environments over time (such as penalties for excessive optimism [Smith et al. 2003]), forecasts of control condition participants are consistent with unadjusted means reported for Sedor's (2002) loss-firm/scenario-structure condition participants (\$0.32 and \$0.72, for years 1 and 2, respectively). This suggests that control condition participants' forecasts exhibit optimism, in particular for year 2 (the year for which Sedor found significant effects).

[12] These difficulty ratings differ significantly between experimental conditions, but both means are at the low end of the scale. This may be the result of participants' ability to access the financial information and the conference-call excerpt (via electronic links) while generating counter-explanations. To the extent that analysts in both conditions easily generated the requested number of counter-explanations, this would bias against finding significant differences in analysts' forecasts across experimental conditions.

[13] Inferences are unchanged when we exclude initial forecast increases in EPS from the model.

[14] We perform a median split instead of a mediation analysis because we lack difficulty ratings for participants in the control condition. Control condition participants were not asked to generate counter-explanations and so were not asked to assess the difficulty of doing so. The modal response to the difficulty item was 1 (on a nine-point scale), the median was 2, and we classified responses of less or equal to the median as "lower" difficulty. With the responses clustered at the bottom of the scale, there is limited variability in this measure, which biases against finding results consistent with our hypotheses and leads our cell sizes to be unbalanced.

Tests of Alternative Explanations

We perform supplemental tests to rule out potential alternative explanations for the results. First, to ensure that the finding of no difference between the control and generate many conditions was not due to small sample sizes, we estimate contrasts for 1,000 bootstrap samples and determine the resulting p-values using the percentile method. Each bootstrap sample was generated by sampling (randomly with replacement) 15 observations within each cell of the design (the smallest of the original cell sizes). For both one-year- and two-year-ahead forecast increases in EPS, inferences are identical to those reported for our main tests of H1 and H2.

Second, we confirmed using analysis of variance that there are no significant relationships between experimental condition and perceptions about task realism, perceptions about the extent to which generating reasons why managers' plans could fail is a routine part of the forecasting process, beliefs about managerial competence, beliefs about managerial honesty, or perceptions regarding the importance of providing favorable or accurate forecasts (p-values > 0.27 for all omnibus F-tests). Therefore, differences in these measures do not explain reported results.

Third, although the ineffectiveness of counter-explanation in the generate many condition is consistent with H2, this result could also be consistent with analysts asked to generate 12 counter-explanations producing lower quality counter-explanations than analysts asked to generate two counter-explanations. To address this possibility, we analyze the content of participants' counter-explanations. We asked two individuals with significant accounting and valuation consulting experience to code independently the 226 counter-explanations provided by participants into content categories (e.g., cannibalization of premium wine sales; changing consumer preferences). One individual identified 20 unique content categories, and both individuals coded participants' counter-explanations into those categories. Cohen's Kappa is 0.76, indicating high to very high initial agreement (Stokes et al. 1995). The coders met to reconcile differences. We use the reconciled codings for all analyses.

We calculate the percentage frequency with which counter-explanations from each category were listed by all participants and use this frequency measure as a proxy for counter-explanation quality. For instance, if a category occurred 20 times, then we assign each counter-explanation from that category the percentage frequency of 20/226 = 8.85 percent. Next, for each participant, we calculate the mean percentage frequency for the counter-explanations he or she generated (i.e., averaging percentage frequencies for the two counter-explanations generated by participants in the generate few condition and for the 12 counter-explanations generated by participants in the generate many condition). Means (standard deviations) for participants in the generate few and generate many conditions are 7.16 (2.67) and 6.99 (1.16), respectively. There are no differences in this measure between experimental conditions ($t_{34} = 0.28$, $p = 0.78$, two-tailed) and thus, differences in counter-explanation quality are unlikely to explain our results.[15]

This test is a conservative one, in that analysts in the generate many condition produced six times as many counter-explanations as did analysts in the generate few condition. Thus, the *total* body of evidence considered by analysts in the former condition would likely be larger than the total body of evidence considered by analysts in the latter condition regardless of its *average* quality. Indeed, the counter-explanations provided by participants in the generate many condition covered a mean of 8.9 content categories (ranging from five to 12), while the counter-explanations provided by participants in the generate few condition covered a mean of 2.0 content categories.[16] As a result, the counter-explanation content considered by analysts in the generate many condition exceeded that considered by analysts in the generate few condition.

IV. EXPERIMENT 2

In Experiment 2, we replicate Experiment 1 and provide additional corroborating evidence that Experiment 1 results are caused by differences in the availability of counter-explanations (i.e., the ease or difficulty of their generation) and not by potential differences in counter-explanation content. We designed Experiment 2 to disentangle the content of the counter-explanations from their availability. Specifically, we added two new experimental conditions that mirror the content of the counter-explanations from the generate conditions while ensuring that all counter-explanations (regardless of number) are equally available to participants (Wanke et al. 1996). In particular, we asked participants in the new experimental conditions to read the counter-explanations generated by other participants (in identical form and order) instead of generating counter-explanations themselves.

If Experiment 1 results are due to differences in the content of the counter-explanations, then those same results should be observed in the new experimental conditions: reading few counter-explanations reduces forecast optimism, whereas reading many counter-explanations does not. In contrast, if Experiment 1 results are due to differences in the availability of the counter-explanations (i.e., the ease or difficulty of their generation), then those results should not be mirrored in the new read conditions. The read task ensures that counter-explanations are easily available to these participants, regardless of whether their matched participants generated few or many counter-explanations. Thus, we expect that reading few and reading many counter-explanations will reduce forecast optimism and make the following unique predictions for Experiment 2:

---

[15] Similarly, order effects could influence the results if participants in the generate many condition did not consider all of their counter-explanations when making their EPS forecasts, but instead, focused on only the first or last few counter-explanations. However, we find no significant differences in the average frequencies for the two counter-explanations generated by participants in the generate few condition when compared to either the first two or the last two counter-explanations generated by participants in the generate many condition ($t_{37} = -0.91$, $p = 0.37$ and $t_{37} = 1.21$, $p = 0.23$, two-tailed, for the first two and last two counter-explanations generated, respectively). Therefore, differences in counter-explanation order are unlikely to be driving results.

[16] Participants in the generate many condition provided similar counter-explanations to those provided by participants in the generate few condition, along with additional items. Although participants in the generate few (generate many) condition used a total of 14 (12) of the 20 content categories in generating their two (first two) counter-explanations, seven of these categories overlap between conditions and account for 74 percent of the first two counter-explanations generated. This percentage does not vary significantly between experimental conditions, with a mean 1.38 (1.53) of the two (first two) counter-explanations provided in the generate few (generate many) conditions coming from these seven categories ($t_{37} = 0.75$, $p = 0.46$). The counter-explanations provided by participants in the generate-many condition ultimately covered all 20 content categories.

H3: Analysts who are asked to read *few* counter-explanations of managers' plans to improve future earnings will make earnings forecasts that are *less* optimistic than the earnings forecasts made by analysts who are not asked to generate or read counter-explanations.

H4: Analysts who are asked to read *many* counter-explanations of managers' plans to improve future earnings will make earnings forecasts that are *less* optimistic than the earnings forecasts made by analysts who are not asked to generate or read counter-explanations.

Design and Method

In Experiment 2, we further explore whether the availability of counter-explanations influences whether a counter-explanation task mitigates scenario-induced optimism (i.e., we replicate Experiment 1, H1 and H2), and we examine whether reading another participant's counter-explanations reduces optimism (i.e., we test H3 and H4). We manipulate the independent variables in a (2 × 2) + 1 (control) between-subjects experiment with the task (generate or read counter-explanations) and the number of counter-explanations (few or many) as independent variables. Control condition participants neither generate nor read counter-explanations.

Participants

Participants are 130 second-year students enrolled in a managerial accounting class in a *BusinessWeek* Top 25 M.B.A. program. Participants were recruited via in-class announcements and assigned to experimental conditions using a predetermined randomization plan. We made a $1,000 contribution to the M.B.A. class gift as a token of our appreciation. Approximately 89 percent of participants report having previous personal investment experience. Participants report having taken an average of four university-level accounting classes and three university-level finance classes.[17] M.B.A. students are appropriate participants for this experiment because the effects of availability on judgment are not expected to depend on characteristics specific to analysts (see Peecher and Solomon 2001; Libby et al. 2002). Additionally, we replicate Experiment 1, H1 and H2 as originally demonstrated with analysts before providing additional evidence that the availability of counter-explanations explains results (i.e., testing H3 and H4).

Case Materials and Procedures

The case materials are similar to those used in Experiment 1. One exception is that the net losses and EPS for the current and prior two years are smaller (e.g., current (year 0) EPS is −$0.08 in Experiment 2, instead of the −$0.64 reported in Experiment 1). We reduce the level of the losses reported in Experiment 2 to increase our ability to generalize the results.[18] Another difference from Experiment 1 is the independent variable manipulations discussed below. The experimental procedures are otherwise identical to those used in Experiment 1: each participant was provided with an electronic link to his or her assigned experimental condition and participants completed the study online.

---

[17] One participant provided inappropriate responses to the forecast question (providing percentages instead of dollar amounts) and one participant failed to answer post-task questions. We eliminated these two observations. Eight M.B.A. students who were not enrolled in the managerial accounting class (because they had previous coursework and were taking a shortened program) but were in the same graduating class asked to participate. They are included in the total of 130. All reported inferences are identical if their data are omitted. Inferences are also unchanged if data from participants with no prior investment experience are eliminated.

[18] Although this change increases our ability to generalize the results, it may limit participants' forecast optimism. To the extent that scenario-induced optimism is greater for loss firms (Sedor 2002), this change in materials would limit our ability to replicate Experiment 1 results and bias against finding support for the hypotheses.

Independent Variables

We asked participants in experimental conditions either to generate counter-explanations or to read the counter-explanations generated by other participants after they read the conference-call excerpt, but before they made their final forecasts. Control condition participants were not asked to generate or read counter-explanations and made final forecasts immediately after reading the conference-call excerpt. Specifically, participants in the generate few and the generate many conditions received the following additional instruction:

> Before moving on to the next section, please list TWO (TEN) thoughts about how future financial performance might not be as positive as Villa Fiore expects. In other words, how might Villa Fiore's plans to increase sales and profitability fail?

The online materials included either two or ten blank lines for participants' counter-explanations, depending on experimental condition. To address the possibility that Experiment 1 results depended on participants struggling to generate the full number of 12 counter-explanations requested before proceeding, we reduced the number of counter-explanations requested in the generate many condition in Experiment 2 to ten, and the instrument was programmed such that participants were required to input a minimum of two counter-explanations before continuing. Allowing participants in the generate many condition to continue without generating all ten counter-explanations could reduce the experienced difficulty of the task, thus reducing the strength of our manipulation and biasing against finding support for the hypotheses.

Participants in the read few (read many) conditions received the following additional instruction:

> Before moving on to the next section, please review the thoughts listed by another analyst about how future financial performance might not be as positive as Villa Fiore expects. In other words, this analyst considered how Villa Fiore's plans to increase sales and profitability might fail.

Each participant in the read conditions was provided with the set of counter-explanations obtained from a matched participant in the generate conditions. In other words, each read few (read many) participant had a unique set of experimental materials that reproduced, in the same form and order, the counter-explanations created by a matched generate few (generate many) participant. This yoked design ensured that the quality and order of counter-explanations were identical between the generate and read few conditions as well as between the generate and read many conditions. However, the matching procedure necessitated collecting data in stages. Our initial invitations to participate were distributed to only a portion of the participants pre-assigned to the control and generate conditions. Subsequent invitations included participants pre-assigned to all five conditions. One consequence of collecting data in stages is that we were unable to ensure equal cell sizes or to match every participant in the generate conditions with a participant in the corresponding read condition. Thus, we exclude from analysis data provided by 13 generate condition participants for whom we did not have matched read condition participants.

Dependent Measures

Participants provided initial and final forecasts of annual EPS for one year and two years ahead. Consistent with Experiment 1, we subtract current (year 0) EPS of −\$0.08

from participants' EPS forecasts to arrive at participants' forecast increases in EPS. The dependent measures are participants' final forecast increases in EPS for both one year (year 1) and two years ahead (year 2). We use participants' initial forecast increases in EPS as covariates in the analyses. Table 2, Panel A reports descriptive statistics for Experiment 2, including means and least squares (adjusted) means for final forecast increases in EPS for years 1 and 2.

Results

Preliminary Analyses and Manipulation Check

We exclude three participants' forecast increases in EPS from the analyses because they differed from sample means by more than four standard deviations, and diagnostic procedures indicated that they are influential outliers. We also exclude the data provided by the three participants matched to these participants to preserve the yoked experimental design. Reported analyses use the remaining 111 observations.

Participants in the generate many condition listed significantly more counter-explanations (i.e., reasons why managers' plans might fail) than did participants in the generate few condition (means = 7.09 and 2.00 for the generate many and few conditions, respectively; $t_{21} = 8.21$, $p < 0.01$). Participants asked to generate ten counter-explanations also assigned marginally higher difficulty ratings to the task than did those participants asked to generate two counter-explanations (means [standard deviations] = 4.72 [2.10] and 3.85 [1.95] for the generate many and generate few conditions, respectively; $t_{40} = 1.40$, $p = 0.08$).

Tests of Hypotheses

Table 2, Panel B reports results of the planned contrasts used to test our hypotheses. Consistent with H1, participants asked to generate few counter-explanations gave marginally less optimistic forecast increases in EPS than did control condition participants for year 1 ($t_{105} = 1.55$, $p = 0.06$) and gave significantly less optimistic forecast increases in EPS than did control condition participants for year 2 ($t_{105} = 2.01$, $p = 0.02$). Consistent with H2, participants asked to generate many counter-explanations gave forecast increases in EPS that did not differ from those of control condition participants ($t_{105} = 0.62$, $p = 0.54$ and $t_{105} = 1.05$, $p = 0.30$ for year 1 and year 2, respectively).[19]

We repeat the main analyses using a median split of the difficulty of generation measure (manipulation check) as the independent variable to corroborate the results. Participants reporting lower difficulty (i.e., 4 or less on a nine-point scale) gave significantly less optimistic forecast increases in EPS than did control condition participants for both year 1 and year 2 (adjusted means = 0.16 and 0.10 for control and lower difficulty, $t_{65} = 2.02$, $p = 0.02$; adjusted means for year 2 = 0.26 and 0.18 for control and lower difficulty, $t_{65} = 1.78$, $p = 0.04$), supporting H1. Participants reporting higher difficulty gave forecast increases in EPS that did not differ from those of control condition participants (adjusted means for year 1 = 0.16 and 0.15 for control and higher difficulty, $t_{65} = 0.35$, $p = 0.73$; adjusted means for year 2 = 0.26 and 0.21 for control and higher difficulty, $t_{65} = 1.17$, $p = 0.25$), supporting H2.

Next, we test the unique predictions for Experiment 2—that participants who read both few (H3) and many (H4) counter-explanations will make less optimistic EPS forecasts than will control condition participants. Both hypotheses are partially supported for year 1 and

---

[19] Applying the bootstrapping technique as described for Experiment 1 improves results slightly in that the H1 contrast is significant for both years 1 and 2 ($p < 0.05$), while the H2 contrast remains insignificant ($p > 0.10$). Inferences for H3 and H4 are identical to those reported below.

TABLE 2
Experiment 2: M.B.A. Students' EPS Forecasts

Panel A: Descriptive Statistics: Mean [standard deviation]

		Year 1			Year 2		
Condition	**n**	**Initial Forecast Increase in EPS**	**Final Forecast Increase in EPS**	**ANCOVA Adjusted Final Forecast Increase in EPS**[a]	**Initial Forecast Increase in EPS**	**Final Forecast Increase in EPS**	**ANCOVA Adjusted Final Forecast Increase in EPS**[a]
Control	27	0.15 [0.13]	0.14 [0.16]	0.15	0.27 [0.23]	0.25 [0.26]	0.25
Generate Few	20	0.23 [0.47]	0.13 [0.27]	0.09	0.34 [0.52]	0.20 [0.30]	0.16
Generate Many	22	0.17 [0.47]	0.13 [0.32]	0.13	0.26 [0.60]	0.20 [0.43]	0.20
Read Few	20	0.10 [0.16]	0.06 [0.13]	0.10	0.18 [0.14]	0.10 [0.15]	0.16
Read Many	22	0.14 [0.11]	0.09 [0.15]	0.10	0.30 [0.24]	0.16 [0.23]	0.14

Panel B: Planned Contrast Tests of Hypotheses

Hypothesis Contrast	**Year 1 t-statistic**	**Year 1 p-value**	**Year 2 t-statistic**	**Year 2 p-value**
H1: Lower relative optimism for generate few versus control	1.55	0.06	2.01	0.02
H2: No difference in relative optimism for generate many versus control	0.62	0.54	1.05	0.30
H3: Lower relative optimism for read few versus control	1.41	0.08	1.98	0.03
H4: Lower relative optimism for read many versus control	1.51	0.07	2.37	0.01

---

[a] The final forecast increase in EPS is adjusted statistically for the initial forecast increase in EPS by including this as a covariate in the ANCOVA model. The adjusted means are reported here.

Panel A provides descriptive statistics by experimental condition. "Year t final forecast increase in EPS" is our dependent variable and is defined as the final (post-manipulation) forecast for year t ($t = 1$ or 2) less the year 0 EPS of −$0.08 as provided to participants. In our analyses, we adjust for the influence of participants' initial (pre-manipulation) forecasts using analysis of covariance (ANCOVA) and thus, we report the ANCOVA least squares adjusted means in Panel A. Participants made initial EPS forecasts after reading summarized financial information and made final EPS forecasts after reading a conference-call excerpt (which presented information about managers' plans to improve future earnings in scenario structure) and generating counter-explanations (if applicable). Participants in the generate few (generate many) condition were asked to list 2 (10) reasons why managers' plans could fail before making final forecasts. Participants in the read few (read many) condition were asked to read the reasons why managers' plans might fail that were produced by matched participants in the generate few (generate many) condition. Participants in the control condition were not asked to generate or read counter-explanations before making final forecasts.

Panel B provides tests of hypotheses comparing the forecast increases in EPS of participants in experimental conditions to those of participants not asked to generate or read counter-explanations (i.e., participants in the control condition) to examine (1) whether the availability of counter-explanations (i.e., the ease or difficulty experienced during the counter-explanation task) affects the relative amount of scenario-induced optimism (H1 and H2) and (2) whether the observed effect is driven by availability and not by variations in counter-explanation content between experimental conditions (H3 and H4). If the forecast increases in EPS in experimental conditions are less than the forecast increases in EPS in the control condition, then experimental condition participants are *relatively less optimistic*. P-values are one-tailed for directional predictions (H1, H3, and H4) and two-tailed for H2.

supported for year 2. Participants asked to read few counter-explanations gave marginally less optimistic forecast increases in EPS than did control condition participants for year 1 (t_{105} = 1.41, p = 0.08) and gave significantly less optimistic forecast increases in EPS than did control condition participants for year 2 (t_{105} = 1.98, p = 0.03). Likewise, participants asked to read many counter-explanations gave marginally less optimistic forecast increases in EPS than did control condition participants for year 1 (t_{105} = 1.51, p = 0.07) and gave significantly less optimistic forecast increases in EPS than did control condition participants for year 2 (t_{105} = 2.37, p = 0.01). Taken together, results support that the availability of counter-explanations is an important determinant of whether counter-explanation reduces scenario-induced forecast optimism.

V. CONCLUSION

In this paper, we examine whether counter-explanation can be used to reduce scenario-induced optimism in analysts' forecasts. In two experiments, we find that the EPS forecasts of professional analysts (Experiment 1) and M.B.A. students (Experiment 2) are influenced by the availability of (i.e., the ease of generating) counter-explanations. Consistent with availability, participants in conditions in which generating counter-explanations was relatively easy made lower EPS forecasts than did control condition participants. Participants in conditions in which generating counter-explanations was relatively difficult made EPS forecasts that did not differ from the forecasts made by control condition participants. In Experiment 1, we perform supplemental tests to rule out alternative explanations and thus, provide evidence suggesting that differences in the ability of counter-explanation tasks to reduce forecast optimism are caused by differences in the availability of counter-explanations.

In Experiment 2, we provide additional evidence that results are consistent with availability influencing counter-explanation's effectiveness. We provided some participants with the counter-explanations generated by other participants (in identical form and order) instead of having these participants generate counter-explanations themselves. Because reading many counter-explanations is at least as effective in reducing optimism as is reading few counter-explanations, we infer that it is the availability of (i.e., the ease of accessing) counter-explanations that drives the reduction in relative optimism observed in Experiments 1 and 2.

This research is the first that we know of to investigate counter-explanation as a means of reducing cognitive biases associated with predictions made by experts. Because managers are likely to continue using scenarios to describe their plans to improve future firm performance, analysts and others who rely on this information when developing forecasts, but wish to avoid optimism, could use a simple counter-explanation task to avoid optimistic predictions.

Our results also suggest that analysts do not appear to consider spontaneously a small number of counter-explanations when initially contemplating managers' scenarios. To the extent that scenario-induced optimism is manifest in analysts' forecasts of annual earnings, analysts' failure to consider counter-explanations of their own accord may be one factor sustaining optimism.

Finally, we demonstrate a boundary condition for the use of counter-explanation as a debiasing mechanism. In both experiments, we find that the availability of counter-explanations (i.e., the ease or difficulty of their generation) influences the ability of this mechanism to reduce scenario-induced optimism in a prediction task. Results suggest that when implementing counter-explanation as a debiasing mechanism, it is important to be aware that a lack of available counter-explanations can dampen the mechanism's effectiveness. A limitation of the study is that using a hypothetical firm precludes us from calculating an absolute measure of forecast optimism and thus, we are unable to determine whether counter-explanation completely eliminates forecast optimism.

Optimism in analysts' long-term forecasts likely is costly to both investors and securities firms. Prior empirical research suggests that correcting for analysts' forecast optimism can improve value estimates (Frankel and Lee 1998). Recently, a number of Wall Street firms agreed to a $1.4 billion settlement for charges by the SEC, NASD, and NYSE, among others, that their analysts routinely issued overly optimistic stock research to investors in order to curry favor with corporate clients and to win investment banking business (Smith et al. 2003). Although these charges presumably relate to forecast optimism driven by economic incentives, addressing more subtle sources of forecast optimism is likely to become of increasing concern to regulators, investors, and Wall Street analysts. Our study introduces a theory-based, yet practical way of reducing scenario-induced optimism in analysts' annual earnings forecasts.

REFERENCES

Abarbanell, J. S., and V. L. Bernard. 1992. Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *The Journal of Finance* 67 (3): 1181-1207.

———, and R. Lehavy. 2003. Can stock recommendations predict earnings management and analysts' earnings forecast errors? *Journal of Accounting Research* 41 (1): 1–31.

Affleck-Graves, J., L. R. Davis, and R. R. Mendenhall. 1990. Forecasts of earnings per share: Possible sources of analyst superiority and bias. *Contemporary Accounting Research* 6 (2): 501–517.

Ali, A., A. Klein, and J. Rosenfeld. 1992. Analysts' use of information about permanent and transitory earnings components in forecasting annual EPS. *The Accounting Review* 67 (1): 183–198.

———, L. Hwang, and M. A. Trombley. 2003. Residual-income-based valuation predicts future stock returns: Evidence on mispricing vs. risk explanations. *The Accounting Review* 78 (2): 377–396.

Arkes, H. R. 1991. Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin* 110 (3): 486–498.

Bonner, S., R. Libby, and M. W. Nelson. 1996. Using decision aids to improve auditors' conditional probability judgments. *The Accounting Review* 71 (2): 221–240.

Bradshaw, M. T., S. A. Richardson, and R. G. Sloan. 2003. Pump and dump: An empirical analysis of the relation between corporate financing activities and sell-side analyst research. Working paper, Harvard University, University of Pennsylvania, and University of Michigan.

———. 2004. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79 (1): 25–50.

Brown, L. D. 1997. Analyst forecasting errors: Additional evidence. *Financial Analysts Journal* (November/December): 81–88.

———. 2001. A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research* 39 (2): 221–241.

Darrough, M. N., and T. Russell. 2002. A positive model of earnings forecasts: Top down versus bottom up. *Journal of Business* 75 (1): 127–152.

DeBondt, W. F. M., and R. H. Thaler. 1990. Stock market volatility: Do security analysts overreact? *American Economic Review* 80 (2): 52–57.

Dechow, P. M., A. P. Hutton, and R. G. Sloan. 1999. An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics* 26: 1–34.

———, ———, and ———. 2000. The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research* 17 (1-3): 1-32.

Dowen, R. J. 1996. Analyst reaction to negative earnings for large well-known firms. *The Journal of Portfolio Management* 23 (1): 49–55.

Dugar, A., and S. Nathan. 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. *Contemporary Accounting Research* 12 (1): 131–160.

Duru, A., and D. M. Reeb. 2002. International diversification and analysts' forecast accuracy and bias. *The Accounting Review* 77 (2): 415–433.

Eames, M., S. M. Glover, and J. Kennedy. 2002. The association between trading recommendations and broker-analysts' earnings forecasts. *Journal of Accounting Research* 40 (1): 85–104.

———, and ———. 2003. Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review* 78 (3): 707–724.

Elgers, P. T., and M. H. Lo. 1994. Reductions in analysts' annual earnings forecast errors using information in prior earnings and security returns. *Journal of Accounting Research* 32 (2): 290–303.

Francis, J., and D. Philbrick. 1993. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research* 31 (2): 216–230.

Frankel, R., and C. M. C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics* 25 (3): 283–319.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40 (1–3): 3–73.

Heiman, V. B. 1990. Auditors' assessments of the likelihood of error explanations in analytical review. *The Accounting Review* 65 (4): 875–890.

Hoch, S. J. 1984. Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10 (4): 649–662.

Hong, H., and J. D. Kubik. 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance* 58 (1): 313–351.

Hunton, J. E., and R. A. McEwen. 1997. An assessment of the relation between analysts' earnings forecast accuracy, motivational incentives and cognitive information search strategy. *The Accounting Review* 72 (4): 497–515.

Irvine, P. 2004. Analysts' forecasts and brokerage-firm trading. *The Accounting Review* 79 (1): 125–149.

Jacob, J., S. Rock, and D. P. Weber. 2003. Do analysts at independent research firms make better earnings forecasts? Working paper, University of Colorado at Boulder.

Kahneman, D., and A. Tversky. 1982. The simulation heuristic. In *Judgment Under Uncertainty: Heuristics and Biases*, edited by D. Kahneman, P. Slovic, and A. Tversky. Cambridge, U.K.: Cambridge University Press.

———, and D. Lovallo. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* 39 (1): 17–31.

Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31 (2): 231–245.

———. 1995. Debiasing the curse of knowledge in audit judgment. *The Accounting Review* 70 (2): 249–273.

Kim, C., and S. Lustgarten. 1998. Broker-analysts' trade-boosting incentive and their earnings forecast bias. Working paper, Queens College of The City University of New York and Baruch College of The City University of New York.

Klein, A. 1990. A direct test of the cognitive bias theory of share price reversals. *Journal of Accounting and Economics* 13 (2): 155–166.

Koonce, L. 1992. Explanation and counter-explanation during audit analytical review. *The Accounting Review* 67 (1): 59–76.

Krishnan, M., and K. Sivaramakrishnan. 1998. A theory of analysts' forecast bias. Working paper, Rutgers University and Texas A&M University.

Lee, C. M. C., J. Myers, and B. Swaminathan. 1999. What is the intrinsic value of the Dow? *The Journal of Finance* 54 (5): 1693–1741.

Libby, R., R. Bloomfield, and M. W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (8): 775–810.

Lim, T. 2001. Rationality and analysts' forecast bias. *The Journal of Finance* 56 (1): 369–385.

Lin, H., and M. F. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25 (1): 101–127.

Liu, J., and J. Thomas. 2000. Stock returns and accounting earnings. *Journal of Accounting Research* 38 (1): 71–101.

Loh, R., and G. M. Mian. 2005. Do accurate earnings forecasts facilitate superior investment recommendations? *Journal of Financial Economics* (forthcoming).

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

McNichols, M., and P. C. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35 (Supplement): 167–199.

Michaely, R., and K. L. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12 (4): 653–686.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* 74 (2): 185–200.

National Investor Relations Institute (NIRI). 2003. NIRI issues survey results on earnings guidance practices; First Call eliminates the term "consensus." *Executive Alert: Updates on Issues Vital to the Practice of Investor Relations.* Available at: http://www.niri.org/irresource_pubs/alerts/ea20031210.pdf.

Nelson's Directory of Investment Research. 2003. Port Chester, NY: W.R. Nelson & Co.

Ohlson, J. A., and B. E. Juettner-Nauroth. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10 (2–3): 349–365.

Peecher, M. E., and I. Solomon. 2001. Theory and experimentation in studies of audit judgments and decisions: Avoiding common research traps. *International Journal of Auditing* 5 (3): 193–203.

Richardson, S., S. H. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Sanna, L. J., N. Schwarz, and S. L. Stocker. 2002. When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory and Cognition* 28 (3): 497–502.

Schwarz, N., F. Strack, H. Bless, G. Klumpp, H. Rittenauer-Schatka, and A. Simons. 1991. Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology* 61 (2): 195–202.

———. 1998. Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment. *Personality and Social Psychology Review* 2 (2): 87–99.

Sedor, L. 2002. An explanation for unintentional optimism in analysts' earnings forecasts. *The Accounting Review* 77 (4): 731–753.

Smith, R., S. Craig, and D. Solomon. 2003. Wall Street firms settle charges over research in $1.4 billion pact. *Wall Street Journal* (April 29): A1.

Sougiannis, T., and T. Yaekura. 2001. The accuracy and bias of equity values inferred from analysts' earnings forecasts. *Journal of Accounting, Auditing & Finance* 16 (4): 331–362.

Stokes, M. E., C. S. Davis, and G. G. Koch. 1995. *Categorical Data Analysis Using the SAS System.* Cary, NC: SAS Institute Inc.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5 (2): 207–232.

Wall Street Journal Online. 2002. Coca-Cola to end practice of giving earnings forecasts. *A Wall Street Journal Online News Roundup* (December 13).

Wanke, M., H. Bless, and B. Biller. 1996. Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin* 22 (11): 1105–1113.

———, G. Bohner, and A. Jurkowitsch. 1997. There are many reasons to drive a BMW: Does imagined ease of argument generation influence attitudes? *Journal of Consumer Research* 24 (2): 170–177.

Zuckerman, G. 2005. CEOs turn mum about projecting earnings. *Wall Street Journal* (March 1): C1, C4.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 399–420

Expense Misreporting in Nonprofit Organizations

Ranjani Krishnan
Michigan State University

Michelle H. Yetman
University of California, Davis

Robert J. Yetman
University of California, Davis

ABSTRACT: We examine whether nonprofit organizations understate fundraising expenses in their publicly available financial statements. A large body of anecdotal evidence notes that an inexplicable number of nonprofits report zero fundraising expenses. We provide empirical evidence that the zero fundraising expense phenomenon is at least partly due to inappropriate reporting. We then examine to what extent these misreported expenses are the result of managerial incentives. Prior research finds an association between reported expenses and managerial compensation as well as the level of donations received. Using these findings we construct two incentive variables and find a positive association between misreporting behavior and managerial incentives. Our results also suggest that the use of an outside accountant reduces the probability that a nonprofit will misreport expenses, consistent with the use of an outside paid accountant increasing the reliability and usefulness of nonprofit financial reports. Finally, we find that SOP 98-2 reduced the probability that a nonprofit will misreport fundraising expenses.

Keywords: *nonprofit organizations; cost shifting; agency problems.*

JEL Classification: *M4; L3.*

I. INTRODUCTION

We examine whether nonprofit managers incorrectly classify costs across the categories of functional expenses reported on their IRS Form 990 (hereafter, IRS 990), and whether this incorrect classification is driven by managerial incentives.[1]

---

[1] The IRS 990 is the primary source of publicly available nonprofit financial information (Joint Committee on Taxation 2000).

We thank two anonymous reviewers, Ashiq Ali, Leslie Eldenburg, Teresa Gordon, Angela Gore, Janet Greenlee, Chris Jones, Lil Mills, Mike Shields, workshop participants at the 2003 American Accounting Association meetings, The University of Arizona, The George Washington University, Michigan State University, and Arizona State University for their valuable comments and advice.

Editor's note: This paper was accepted by Terry Shevlin, Senior Editor.

Submitted June 2005
Accepted September 2005

The IRS 990 partitions expenses into programs (i.e., charitable), fundraising, and administrative categories. Prior research provides evidence that two metrics based on these three expense classifications are of particular importance because they influence managerial compensation as well as the decisions of donors and other stakeholders (Baber et al. 2002; Okten and Weisbrod 2000; Posnett and Sandler 1989; Weisbrod and Dominguez 1986). The first metric is the program ratio, which is the amount of program expenses divided by total expenses. Higher program ratios indicate that a larger proportion of total expenses are directed toward the charitable mission. Various charity watchdog agencies as well as state regulators suggest that higher program ratios are indicative of more efficient and effective organizations. The second metric is the fundraising ratio, which is the amount of fundraising expenses divided by total donations. Lower fundraising ratios are suggestive of higher fundraising efficiency because fewer donations are consumed by the cost of fundraising activities.

Concerns over the quality of nonprofit financial reporting are increasing (Yetman and Yetman 2004). Congress recently commissioned the Government Accountability Office (GAO) to conduct a study of nonprofit reporting (United States GAO 2002). The GAO found several reporting weaknesses, including finding that over one-half of all nonprofits that earn donations fail to report any fundraising expenses. Although the GAO report and many subsequent informal analyses (e.g., Froelich et al. 2000; Hager 2003) conclude that this finding is curious and inexplicable, scant empirical evidence exists that this phenomenon is the result of incorrect reporting.

We conduct two separate empirical tests to determine whether nonprofits inappropriately understate their fundraising expenses and overstate their program expenses. For the first test, we randomly select a sample of nonprofits that receive donations and report zero fundraising expenses on their IRS 990, and search their Internet web pages for evidence that they conduct fundraising activities. We find that nearly half of the websites we examined show evidence of fundraising activities of the type that would give rise to reportable fundraising expenses that ostensibly should have been reported on the IRS 990. For the second test we compare a sample of audited financial statements to matched sets of IRS 990s. We find that 38 percent of the observations that reported some fundraising expenses on their audited financial statements reported lower (and in many cases zero) amounts of fundraising on their IRS 990s, and that the mean amount of fundraising reported on the 990 is significantly less than that reported on the financial statements.

We next examine the firm-specific characteristics that are associated with these misallocated expenses. The primary goal of this analysis is to determine whether nonprofit expense misreporting is random error, related to managerial incentives, or the result of other factors such as lack of expertise. We examine two sources of incentives for managers to misallocate expenses: to garner higher managerial compensation and to attract additional donations revenues. We compute firm-level incentive variables using time-series associations between the program ratio and managers' compensation and donations revenue. We find our incentive measures to be positively associated with the incidence of zero fundraising expenses. We also find that the use of an outside accountant and accounting rule changes reduces the incidence of zero fundraising expenses. Our results are robust to controlling for expertise (as measured by size and age of the charity).

As previously noted, a growing body of anecdotal evidence points out that an inexplicable number of nonprofits report zero fundraising expenses. Our paper makes a contribution by being among the first to provide empirical evidence that this phenomenon is due to incorrect reporting. A recent paper by Jones and Roberts (2006) examines if nonprofits misallocate the costs of combined fundraising and educational campaigns (so called "joint costs"). They find that nonprofits over-allocate expenses to education and under-allocate expenses to fundraising, thus simultaneously improving their program and fundraising ratios. Although Jones and Roberts (2006) examine a different setting (i.e., the subset of expenses classified as "joint costs") using a different sample (i.e., only those observations that report "joint costs," which is about 2 percent of the nonprofit population), their results are consistent with ours and provide additional evidence that nonprofit organizations misreport their financial reports.

In addition to documenting expense misallocations by nonprofits, our paper also provides evidence that nonprofit managers respond to incentives to manage their publicly available financial information by misstating the amounts of expenses, which is important in that it addresses the issue of whether these misallocations are due to random error, intentional manipulation, or some other factor.

Our results have important implications for nonprofit researchers as they suggest that organizations that report zero fundraising expenses should be carefully scrutinized. Depending on the type of analysis and question being asked, it might be prudent for a researcher to exclude organizations that report zero fundraising expenses from their analysis. In addition, our study has important policy implications. From a public policy perspective, to the extent donors use the disclosed figures for making their donations allocation decisions, resources are misallocated among nonprofits. Also, our study provides potentially useful information to governmental regulators responsible for nonprofit financial information accuracy, especially in light of the recent attention given to nonprofits' reported fundraising activities by the courts (e.g., in May, 2003, the United States Supreme Court ruled that states can pursue fraud charges against fundraisers who intentionally mislead donors; see 538 U.S., 2003 *Madigan v. Telemarketing Associates*) and by the previously mentioned GAO report.

The next section discusses the agency problem in nonprofit organizations and its influence on nonprofit financial reporting. Section III provides evidence that nonprofits understate their fundraising expenses and thereby likely overstate their program expenses. In Section IV we examine the characteristics that are associated with the incidence of zero reported fundraising expenses and attempt to determine if observed behavior is consistent with managerial incentives. The final section concludes and discusses the implications of our findings.

II. AGENCY PROBLEMS IN NONPROFIT ORGANIZATIONS

Because contracts usually cannot be costlessly written and enforced, most organizations, regardless of their ownership structure, face agency problems (Fama and Jensen 1983; Jensen and Meckling 1976). In publicly traded firms, the agency problem arises because operating decisions are made by managers whose interests are not necessarily identical to those of the shareholders. Economists suggest that in some types of services, the nonprofit organizational form may help reduce these agency conflicts (Arrow 1963; Fama and Jensen 1983; Hansmann 1996). This literature argues that nonprofit organizations arise in areas such as healthcare, in response to the asymmetry of information between the consumers and the producers of a service. To instill trust in the recipient, the providing firm uses the nonprofit form of organization to provide a signal that the firm is not interested in profit maximization, but rather is working to maximize the recipient's welfare. Using similar logic, Fama and Jensen (1983) conclude that, "the absence of residual claims avoids the donor-residual claimant agency problem and explains the dominance of nonprofits in donor-financed activities."

However, the absence of residual claimants does not automatically ensure that nonprofit organizations will be free of agency problems. Managers can consume more resources than is required for efficient production for various reasons. For example, managers can consume excessive perquisites, exert lower effort, or expend resources on activities peripheral to the stated mission of the organization, leading to agency losses. If managers are not effectively monitored by donors, then the potential for agency losses increases (Hansmann 1996). Under these circumstances, accounting information can assist donors and other stakeholders in monitoring their implicit contracts by providing a means for donors to evaluate whether the nonprofit is using its donations in the most efficient and effective manner.

Indeed a substantial body of research shows that donors use accounting measures to monitor the efficiency of a nonprofit organization to ensure stewardship of their donated resources and to assist them in making their donation allocation decisions (Baber et al. 2001; Callen 1994; Okten and Weisbrod 2000; Posnett and Sandler 1989; Tinkelman 1999; Weisbrod and Dominguez 1986; Yetman and Yetman 2003). This literature concludes that donors consider the program and fundraising ratios as important measures of nonprofit organizational efficiency and empirically documents that donations increase as program ratios increase (Harvey and McCrohan 1988; Khumawala and Gordon 1997; Weisbrod and Dominguez 1986) and decrease as fundraising ratios increase (Okten and Weisbrod 2000).

Because program ratios are an important factor for donors, compensating a manager based on program service ratios is one way to increase goal congruence between the donors and the managers such that when the managers maximize their utility, donors' utility is also maximized, thereby reducing agency losses. Consistent with this logic, prior research has documented a positive relationship between program ratios and compensation (Baber et al. 2002).

When accounting information is used as a basis for managerial compensation and donation decisions, one implication that follows is that managers would prefer to disclose accounting information that makes it appear as though the nonprofit is functioning efficiently. A number of prior accounting studies in the *for-profit* setting including Healy (1985), Holthausen et al. (1995), and Guidry et al. (1999) find that managers who have the economic incentive (such as bonus tied to financial performance) manipulate earnings to increase their compensation. In the following empirical analysis we first determine if *nonprofit* managers misreport expense allocations in their financial reports by shifting costs away from fundraising and toward programs. We then test our agency predictions by examining the association between the strength of managers' incentives and financial statement management.

III. EXPENSE MISREPORTING ON THE IRS FORM 990

As previously noted, a large number of nonprofits report zero fundraising expenses. Anecdotal literature speculates that this behavior is inappropriate and perhaps even intentional, as a nonprofit can improve its program ratio by shifting costs from fundraising expenses to program expenses. The purpose of this portion of our analysis is to determine whether the widespread anecdotal speculation can be supported by empirical evidence that nonprofits are inappropriately presenting their financial results to the public.

We address the issue by investigating expense allocations on IRS 990, which is the principle source of nonprofit financial information. All organizations granted federal tax-exemption with gross receipts of more than \$25,000 must file a 990.[2] Penalties, including imprisonment and loss of tax-exempt status, apply to fraudulently filed IRS 990s, although audit rates are low. Affiliated or related organizations are also required to file a complete 990 unless they choose to be included in a combined return.

IRS 990 data are available from two sources, and we use both of these databases for various portions of our analyses. The first is the National Center for Charitable Statistics (NCCS) database, which includes computer readable data for the years 1982 through 2001 (exclusive of 1984 when no sample was taken). The NCCS database includes all nonprofits with total assets greater than \$25 million, plus a stratified random sample of smaller nonprofits. Thus, the NCCS database does not include all nonprofits in every year, but it does capture the bulk of the economic activity of the nonprofit sector. The second nonprofit database is from Guidestar, a project of the Urban Institute sponsored by the Internal Revenue Service. Guidestar provides scanned images of IRS form 990s on their web page (http://www.guidestar.org) and is the most comprehensive database of nonprofit financial information available, but unlike the NCCS data discussed above, the Guidestar data must be hand collected.

We restrict our samples to IRC 501(c) (3) organizations but exclude private foundations.[3] We compare reported amounts of fundraising and program expenses on the IRS 990 to two other distinct sources of data including Internet web pages and financial statements. We presume that audited financial statements are less likely to be subject to manipulation and are therefore useful to calibrate the accuracy of the reported amounts on the IRS 990; we discuss this issue in more detail below.

Comparison of the IRS 990 to Internet Web Pages

For our first test we use the samples of educational, medical, and charitable nonprofit IRS 990 data as contained in the 1998 NCCS database. We use 1998 data as it was the most recent year of data available when this portion of our study was conducted. We limit our analysis to those nonprofits that receive public donations (from individuals and corporations) because the other two types of donations (i.e., government grants and feeder donations such as those from the United Way) can frequently be raised with little consistent fundraising effort.[4] The intent of our analysis is to examine a sample of nonprofits that receive public donations and yet report zero fundraising expenses to determine if they should be reporting non-zero amounts of fundraising expenses. We present the results of this analysis in Table 1, which shows that in the 1998 sample there were 663 educational, 2,090 medical, and 2,242 charitable nonprofits that reported zero fundraising expenses

---

[2] Foreign nonprofits cannot solicit tax-deductible contributions in the United States, nor are they generally granted tax exemption on their activities within the U.S. All nonprofits granted federal and state tax exemption must be incorporated in a U.S. state (or Puerto Rico or the District of Columbia) and thus would be subject to form 990 requirements.

[3] Many 501(c)(3) public charities use the word "foundation" in their title even though they are legally organized and operated as public charities. See Sansing and Yetman (2006) for a discussion of the differences between public charities and private foundations.

[4] Feeder organizations undertake fundraising activities and then distribute donations received to downstream nonprofits. Although downstream nonprofits have to incur at least some amount of fundraising to obtain feeder funding, it is possible that the amounts expended per dollar of donations are small. With respect to government grants, fundraising can be lumped into one year, while a multiyear grant could be received over several years with no additional fundraising effort.

TABLE 1
Internet Sample

Panel A: Descriptive Statistics

Variable	Mean	First Quartile	Median	Third Quartile	Standard Deviation
Total Revenue (in $thousands)	68,585	2,159	11,052	44,040	251,525
Public Donations (in $thousands)	4,260	242	659	1,884	17,066
Total Assets (in $thousands)	98,967	6,665	30,399	78,282	398,659
Program Ratio	0.82	0.77	0.87	0.94	0.20
Age	29	13	26	46	19

Panel B: Internet Search Analysis of Nonprofits that Report Zero Fundraising Expenses on their IRS 990

	Sample Frame	Percent Reporting Zero Fundraising	Average Public Donations (in $thousands)	Websites (%) Searched	Websites Found	Websites with Evidence of Fundraising	Number with Related Fundraising Foundations	Number (%) of Remaining Websites with Evidence of Fundraising
Educational	663	25%	3,650	66 (10%)	38	26	3	23 (60%)
Medical	2,090	46%	1,385	209 (10%)	91	13	7	6 (7%)
Charitable	2,242	33%	1,751	224 (10%)	125	72	0	72 (58%)

This table examines the Internet sample, which is a randomly chosen sample of 10 percent of the nonprofit organizations from the 1998 NCCS database, which report the receipt of public donations from individuals or corporations, but report zero fundraising expenses. We considered evidence of fundraising activities to be explicit requests for donations including providing instructions to individual donors on how to give *and* the contact name of at least one person in charge of fundraising activities.

although they received public donations.[5] We randomly selected (using a random number generator) 10 percent of these zero reporters (i.e., 66 educational, 209 medical, and 224 charitable nonprofits) for further examination.[6] Panel A of Table 1 provides descriptive statistics, which shows that the Internet sample has average revenues of $69 million, public donations of $4 million, and total assets of $99 million in 1998. In addition, the average program ratio is 0.82 and the average age is 29 years.

Next we searched for our sample organizations' Internet websites using the Yahoo! and Google search engines and were able to find Internet websites for 38 educational, 91 medical, and 125 charitable nonprofits. Two investigators (one of the authors and a graduate research assistant) conducted separate extensive detailed searches of each Internet website for evidence of fundraising activities that would produce reportable expenses on the IRS 990s. For our purposes, we consider evidence of reportable fundraising activities to be direct solicitations or instructions on how a potential donor could contribute *and* direct evidence of at least one individual responsible for fundraising activities (such as an employee with the title of "director of fundraising" or "director of donations"). We classified observations as having evidence of fundraising activities only if both investigators found both types of evidence (i.e., solicitation and personnel) on the website. During our Internet website search we found evidence of fundraising activities that met our hurdle for 26 educational, 13 medical, and 72 charitable organizations.

We recognize that some nonprofits have related unconsolidated fundraising foundations that perform their fundraising. To address the effects of fundraising subsidiaries we searched the Guidestar nonprofit database for related entities for the sample of 26 educational, 13 medical, and 72 charitable nonprofits for which we found evidence of fundraising. After removing observations that have related fundraising foundations, we find that 23 of 38 (60 percent) of educational, 6 of 91 (7 percent) of medical, and 72 of 125 (58 percent) of charitable nonprofits are likely to have inaccurately reported their fundraising expenses on their IRS 990. In total we find that about 40 percent of nonprofits examined (101 out of 254) that report zero fundraising expenses on their IRS 990 probably incur some reportable fundraising expenses.

We also recognize that it is possible that the organizations used volunteer labor for their fundraising activities. For this reason our search was for not only a fundraising employee, but also for evidence of solicitation activities. By their nature, solicitation activities are not donated (electricity, computer time, supplies costs, allocated building costs, etc.) and would thus produce reportable fundraising expenses on the IRS 990. However, to gain further assurance that the identified employees were not all voluntary, we examined the IRS 990s of all organizations in the sample for evidence that they used volunteer labor (line 82a on part VII of the IRS 990). We found that only three of our sample organizations reported receiving volunteer labor, and those amounts were relatively small. Thus, we feel confident that the use of volunteer labor is not the cause of zero reported fundraising expenses as identified by our Internet search. In summary, this portion of our analysis provides evidence that nonprofits understate the extent of their fundraising activities on their IRS 990s.

---

[5] Note that, on average, 37 percent of our sample observations do not report fundraising expenses, while approximately 50 percent of the population does not report fundraising. Our sample has a lower percentage of zero reporters because we exclude organizations that earn their donations from feeder organizations or from the government, which as previously discussed can be earned with reduced fundraising efforts.

[6] We limit our analysis to the zero reporters because our intent in this section is to provide evidence that the zero fundraising reporters actually incur some fundraising.

Comparison of the IRS 990 to Audited Financial Statements

For the second test of understated fundraising expenses, we compare a sample of audited financial statements to matched sets of IRS 990s. We presume that the audited financial statement is more accurate with respect to reported amounts of fundraising expenses as the financial statements are required to be prepared in accordance with Generally Accepted Accounting Principles (GAAP) and are subject to independent verification by auditors. Nonprofits are not required by federal law to produce financial statements (other then their IRS 990) nor are they required to have audited any financial statements they do prepare. However, it is common for nonprofits to prepare financial statements and undergo audits for several reasons. First, many granting agencies such as the United Way and federal and state governments frequently require audits as a condition of funding. Second, lenders frequently require audited financial statements as a condition of funding (over 60 percent of nonprofits have some form of debt outstanding). Third, many boards of directors require audits for their nonprofits for financial oversight reasons. Finally, many states require nonprofits with assets above certain threshold amounts to undergo financial statement audits if they solicit donations within that state. Because financial statements are generally required for some specific purpose or request, they typically are not available to the general public. Nonetheless, audited nonprofit financial statements must conform to GAAP and should properly report the amounts expended on programs, administration, and fundraising to the extent those amounts are material.

To obtain a sample of audited financial statements, we sent a request to 1,824 organizations requesting copies of their audited financial statements. Of the 1,824 observations queried 216 (12 percent) voluntarily supplied us with copies of their audited financial statements. On average each nonprofit sent us two years of financials, providing us with a sample of 401 firm-year observations. We excluded observations that did not earn any public donations (i.e., from individuals and corporations) for reasons discussed earlier and observations that reported zero fundraising expenses on their audited financial statements (but we did not exclude observations that reported zero fundraising expenses on their IRS 990).[7] We also took great care to ensure that the financial statements and the related IRS 990s were for the same entities and groups by matching on total revenues, total expenses, addresses, and employer identification numbers. After these screens, 101 firm-year observations representing 55 unique nonprofits remain.[8]

Table 2, Panel A reports descriptive statistics, which shows average revenues of around \$40 million, public donations close to \$14 million, and total assets of \$108 million. The average program ratio is 0.77 and the average age of the nonprofit is 42 years.[9] Table 2, Panel B shows that the mean reported fundraising on the financial statements is larger than that reported on the IRS 990 by about \$2.7 million, which is statistically significant. Of the 101 observations in the sample, 52 organizations report nearly the same fundraising

---

[7] A nonprofit must report fundraising expenses, regardless of the magnitude, on its 990. Thus, if a nonprofit reports some fundraising on its financials but not on its 990s, then it is likely that the nonprofit misreported fundraising expenses on its 990. However, due to materiality considerations that drive financial statement preparation (but not the 990s) it is not possible to draw any inferences when a nonprofit reports some fundraising expenses on its 990 but zero on its financial statements.

[8] We replicate our analysis by year, such that each nonprofit is included only once in each yearly analysis. The results are substantively the same.

[9] In contrast, untabulated analysis based on the population of nonprofits available in the NCCS database reveals that the response sample reports smaller program ratios than the population, which is consistent with nonprofits that are less aggressive misreporters being willing to voluntarily supply their otherwise confidential audited financial statements. If it is true that less aggressive nonprofits were more willing to supply us with their financial statements, then our estimates of understated fundraising expenses on the 990 are conservatively biased.

TABLE 2
Sample of Nonprofits with Matched Sets of Audited Financial Statements and IRS Form 990s

Panel A: Descriptive Statistics

Variable	Mean	First Quartile	Median	Third Quartile	Standard Deviation
Total Revenue (in $thousands)	40,329	7,816	17,293	40,951	76,152
Public Donations (in $thousands)	13,530	1,702	4,105	9,883	37,484
Total Assets (in $thousands)	107,771	14,689	45,200	91,856	200,920
Program Ratio	0.77	0.72	0.81	0.87	0.15
Age	42	32	41	53	18

Panel B: Comparison of Reported Fundraising Expenses on Matched Sets of Audited Financial Statements and IRS Form 990s

	Matched Sample	Equal Fundraising[a] (+/−5%)	Overstated Fundraising in 990[b]	Understated Fundraising in 990[b]	t-Test for [Overstated] = [Understated]	Portion of Firms with Understated Expenses in 990 that Report $0 on 990
Observations	101	52	11	38		14
Fundraising$_{FS}$ – Fundraising$_{990}$:						
Mean difference	2,661,998**	−5,403	−673,960	7,277,797	1.92**	619,066**
Mean difference as a % of total fundraising expenses reported in financial statements	0.22***	0.00	−0.25	0.66	2.92**	1.00***
Program Ratio$_{FS}$ – Program Ratio$_{990}$:						
Mean Difference	0.014	0.00	−0.01	0.07	5.99***	

---

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively.

This table examines a hand-collected sample of 101 observations across 55 unique nonprofits with matched audited financial statements and IRS 990s that reported some amount of fundraising expenses on their audited financial statements.

[a] Firms in the Equal Fundraising column include observations where the reported fundraising on the two forms match to within 5 percent of the amount reported on the audited financial statements.

[b] Overstated (understated) in 990 includes those observations where the amount of fundraising in the IRS 990s is greater (less) by 5 percent of the amount reported in the financial statements.

expenses on the IRS 990s versus the audited statements (i.e., within 5 percent of the amount reported on the financial statements), 11 report more fundraising on their IRS 990s, and 38 report less fundraising on their IRS 990s.[10] For the 11 that report more fundraising on their IRS 990s, the mean difference is $673,960, whereas for the 38 that report less fundraising on their 990, the mean difference is more than ten times greater ($7,277,797); the difference between the two groups is statistically significant. Because these unscaled numbers are likely subject to skewness, we also examine the program ratio and again find a statistically significant difference. For the 11 observations that report more fundraising on their 990, the mean effect on the program ratio is only 1 percent, whereas for the 38 that report less fundraising on the 990, the mean effect is 7 percent; the difference is statistically significant. Interestingly, we find that of the 38 observations that underreported their fundraising on their IRS 990s, only 14 of them reported zero fundraising. Thus, the occurrence of underreported fundraising is pervasive to many nonprofits, including those who report non-zero fundraising in their IRS 990s.

To summarize our results thus far, our study provides evidence that nonprofit organizations are understating their fundraising expenses, and as a result are likely overstating their program expenses. This has the effect of making nonprofits appear more efficient to stakeholders and regulators, potentially affecting resource allocation and regulating decisions. Our results provide among the first empirical evidence supporting the widely held belief that expenses reported on the IRS 990 are likely to be misstated.

IV. ANALYSIS OF COST-SHIFTING INCENTIVES

The purpose of this portion of our analysis is to determine if the incidence of reporting zero fundraising expenses is the result of random error, accounting rule changes, expertise, or managerial incentives consistent with the predictions of agency theory. If zero reporting is random, then we will fail to find any statistically significant association with managerial incentives or other factors. In contrast, if expense misreporting is associated with managerial incentives, then our results have important policy implications (i.e., for donations allocation and governance decisions by stakeholders) as well as implications for researchers who use nonprofit data as the basis of their empirical analyses.

Measuring managerial incentives to manipulate financial information is a difficult task as it necessarily involves measuring an unobservable (latent) variable. Fortunately, prior literature identifies two clear incentives a nonprofit manager has to manipulate their information and provides us with ways to measure these incentives on a firm-specific basis. As discussed earlier, the first incentive arises from nonprofit managers' desire to maximize managerial pay, and the second incentive arises from nonprofit managers' desire to maximize donations received. To estimate these incentives at the firm level, we rely on the findings of prior research that has shown a direct relationship between nonprofit managerial compensation and reported program ratios (Baber et al. 2002). The results of Baber et al. (2002) are consistent with nonprofit firms using compensation to align the incentives of managers and external stakeholders, and to reduce agency losses. Additionally, prior research finds that donors use similar ratios to evaluate the efficiency of nonprofits' operations while making their donation allocation decisions (Baber et al. 2001; Callen 1994; Okten and Weisbrod 2000; Posnett and Sandler 1989; Tinkelman 1999; Weisbrod and Dominguez 1986; Yetman and Yetman 2003).

---

[10] Similarly, Froehlich et al. (2000) find that 50 percent (average of 45 percent in 1988 and 55 percent in 1994) of their sample reports the same amount of fundraising expenses on their audited financials and on their 990 (matched to within 10 percent).

We next discuss the details of how we derive these two incentive metrics. We refer to the metrics as sensitivity variables, as they are intended to capture the sensitivity of compensation of the nonprofit manager to the reported program ratio, and the sensitivity of donations to the reported program ratio.

Incentive Variables

Compensation Sensitivity Variable

Baber et al. (2002) conjecture and find evidence that changes in compensation are significantly associated with accounting-based performance measures, measured by changes in program spending resulting from both changes in revenue raised and changes in the program spending ratio. They separate these components in their model, where *Comp* is compensation paid to CEO, *Revenue* is total revenue, and $\Delta Yield$ measures the extent to which the change in program spending is attributable to changes in the program spending ratio.

Following Baber et al. (2002), we derive our organization-specific compensation sensitivity variable by estimating the following model by firm:

$$\%\Delta Comp_t = \beta_0 + \beta_1 \%\Delta Revenue_t + \beta_2 \%\Delta Program\ Ratio_t + \varepsilon_t, \tag{1}$$

where:

$$\%\Delta Comp_t = \frac{Comp_t - Comp_{t-1}}{Comp_{t-1}},$$

$$\%\Delta Revenue_t = \frac{Revenue_t - Revenue_{t-1}}{Revenue_{t-1}},$$

$$\%\Delta Program\ Ratio_t = \frac{\left(\frac{Program_t}{Revenue_t} - \frac{Program_{t-1}}{Revenue_{t-1}}\right) Revenue_t}{Program_{t-1}},$$

Comp is officer's compensation from line 25a of IRS 990, *Revenue* is total revenues from line 12 of IRS 990, and *Program* is program expenses on line 13 of IRS 990. With the exception of *Comp*, these variable definitions follow those in Baber et al. (2002).[11] Note that our variable *Program Ratio* is identical to *Yield* in Baber et al. (2002).

Using a cross-sectional analysis, Baber et al. (2002) find a statistically significant positive β_2 in Equation (1), which they interpret as evidence that nonprofit officer's compensation is increasing in the reported program ratio. We use a time-series analysis by firm to derive organization-specific estimates of β_2, which measure the sensitivity of compensation to the reported program ratio, after controlling for the effect of changes in revenue. The more sensitive a manager's pay is to the reported program ratio, the more incentive a manager has to improve that ratio so as to garner higher compensation.

---

[11] We measure *Comp* as the total compensation paid to officers, directors, and key employees, whereas Baber et al. (2002) use the salary of either the CEO or the highest paid employee. We do not apply the Baber et al (2002) procedure because about 30 percent of organizations do not report CEO compensation, or provide specifics about the compensation paid to the five highest-paid employees, and our objective is to examine the entire population of nonprofits that file 990s.

Donations Sensitivity Variable

We derive our organization-specific donations incentive variable by estimating the following model by firm:

$$\%\Delta Dons_t = \beta_0 + \beta_1 \%\Delta Revenue_t + \beta_2 \%\Delta Program\ Ratio_t + \varepsilon_t \tag{2}$$

where:

$$\%\Delta Dons_t = \frac{Dons_t - Dons_{t-1}}{Dons_{t-1}}$$

and *Dons* is public donations from line 1a of IRS 990. This model is a straightforward extension of Model 1, where we replace the change in compensation by change in donations as the dependent variable. Although Baber et al. (2002) do not investigate the effects of changes in the program ratio on changes in donations received, there is considerable support in the literature that donations are positively associated with the program ratio (see Okten and Weisbrod [2000] for a discussion of this literature).[12]

The organization-specific coefficients β_2 are estimates of how sensitive changes in donations received are to changes in reported program ratios, after controlling for the effects of changes in revenue. The more sensitive the amount of donations received is to the reported program ratio, the more incentive a manager has to improve that ratio so as to garner higher amounts of donations.

Empirical Estimation of Incentive Variables

We use data from the 1985 to 2001 NCCS database and for reasons discussed earlier require that each firm reports non-zero values for all variables and that they do not report any government grants or feeder donations. We require at least eight usable years of data (results are not sensitive to required time-series from six to 12 years). To control for outliers in Models (1) and (2) we winsorize all observations at the 1st and 99th percentile as well as remove observations with a studentized residual greater than 2 or less than −2. After these screens there are 6,826 valid observations for Model 1 and 26,374 valid observations for Model 2.[13,14]

---

[12] There are many alternative models that we could use to estimate the relationship between donations and the reported ratios other than this straightforward extension of Equation (1). To ensure that our results with respect to the donations sensitivity variable are not driven by the specific model above, we also estimate the sensitivity variable using the methodology of Greenlee and Brown (1999), who examine the effects of the administrative ratio (i.e., the ratio of administrative expenses to total expenses or to total revenues) on donations with no changes in the inferences of our results. A separate line of literature examines the effects of reported fundraising expenses on donations (Okten and Weisbrod 2000; Weisbrod and Dominguez 1986). We are not able to use these models to estimate our sensitivity variables as they did not directly include the program ratio as a regressor, but rather included the ratio of fundraising expenses to total expenses, which is zero for many observations in our sample. The presence of zero values reported by several nonprofits over time would produce unusable sensitivity measures because of our firm-specific regression methodology.

[13] We are unable to conduct this analysis using only the sample of firms that are analyzed in Section III (Table 2), because of sample size limitations. After applying the constraints required for Models 1–3 (such as number of years of firm-level data availability) on our already small samples from Section III, we are left with too few observations to estimate our regression model. Hence, we use the broad set of nonprofits and use an indicator variable for zero reported fundraising expenses as a proxy for underreported fundraising expenses, because our analysis in Section III suggests that these organizations are misreporting their fundraising expenses.

[14] Recall that the data screen for Model 1 required positive values for the change in compensation variable, which was not the case for many observations, causing the number of usable observations for Model 1 to be smaller than those for Model 2. To ensure that our results are not sensitive to this, we replicated our results in Model 2 limiting the analysis to the same observations used in Model 1 (if the data were not missing) with no change in the inferences of our results.

Table 3 presents the descriptive statistics for the regression variables used for estimating Models (1) and (2) as well as the results of those regressions. The mean and median officer's compensation (in total for the nonprofit) are \$407 and \$204 thousand, respectively. Baber et al. (2002) found a mean and median of \$131 and \$113 thousand, respectively. Our numbers are higher as we use the total amount of officer's compensation per nonprofit, which aggregates the pay of key officers, whereas Baber et al. (2002) report the average compensation paid to a single person. The mean and median percentage changes in officer's compensation are 7.66 and 6.97, respectively, which are very similar to those found by Baber et al. (2002) (i.e., mean of 6.84 and median of 5.65). The closeness of our compensation change statistics to those of Baber et al. (2002) suggests that the use of aggregate managerial pay is reasonable. The mean and median changes in the *Program Ratio* are −1.16 and 0.21, which are also reasonably close to the figures obtained by Baber et al. (2002) (i.e., mean of −2.18 and median of −0.40). The similarity of our statistics with those found in Baber et al. (2002) provide comfort that our sample composition, although different from Baber et al. (2002), is reasonable for the given analysis.

In the lower portion of Table 3 we present the results of the firm-specific regressions. Approximately 25 percent of our observations have negative values for their firm-specific estimated sensitivity measures, which we recognize lack economic meaning and are an artifact of the estimation process due to small numbers of observations at the firm level. Because these negative sensitivities make no economic sense, we reset all negative values to be equal to zero because in theory these sensitivity variables are bounded below at zero. We use these truncated sensitivity measures for our primary tests, although we also undertook two alternative approaches to ensure that our results were not driven by this truncation. First, we use the sensitivity variables as produced and, second, we exclude observations with a negative sensitivity. Although the magnitudes of the results vary with these additional tests, the general tenor of the results remains unchanged with respect to statistical inferences.

For the compensation sensitivity regression we find a mean (median) organization-specific coefficient for the change in the program ratio of 0.414 (0.099). Baber et al. (2002) obtained a pooled-cross sectional coefficient estimate of 0.076 for their full sample, which is close to our median estimate. With respect to donations sensitivity, we find mean (median) values of 0.866 (0.076). We hypothesize that the benefits (i.e., payoff) to managing the reported amounts of program and fundraising expenses is increasing in these two sensitivity measures, and therefore ratio management (measured as reporting zero fundraising expenses) will be increasing in these sensitivity metrics. In addition to these two sensitivity variables, we include other independent variables, which we discuss in the next section.

The Association between Managerial Incentives and Reporting Zero Fundraising

The compensation and donation sensitivity measures are intended to capture firm-specific incentives to improve the fundraising and/or program ratios. As previously noted, one way to improve the ratios is to under-report the amount of fundraising expenses and over-report the amount of program expenses. We use under-reported fundraising expenses as our dependent variable to examine whether managers whose compensation and donations are relatively more sensitive to the reported program ratio are more likely to improve their financials by understating their fundraising expenses.

Other Independent Variables

We include two measures reflecting changes in accounting rules. The first is an indictor variable equal to 1 if the year is greater than or equal to 1995, which represents the effective

TABLE 3
Descriptive Information and Regression Results for Financial Statement Management Incentive Variables

$$\%\Delta Comp_t = \beta_0 + \beta_1\%\Delta Revenue_t + \beta_2\%\Delta Program\ Ratio_t + \varepsilon_t$$

$$\%\Delta Dons_t = \beta_0 + \beta_1\%\Delta Revenue_t + \beta_2\%\Delta Program\ Ratio_t + \varepsilon_t$$

Variable	n	Mean	First Quartile	Median	Third Quartile	Standard Deviation
Regression Variables						
Officer's compensation	6,826	406,729	99,000	204,450	442,099	828,529
Percentage change in officer's compensation (*%ΔComp*)	6,826	7.66	1.36	6.97	10.00	20.59
Public Donations (in $thousands)	26,374	3,739	99	536	2,338	13,617
Percentage change in public donations (*%ΔDons*)	26,374	11.51	−27.63	4.06	51.39	55.81
Total revenues (in $thousands)	26,374	30,606	3,666	10,616	30,944	61,539
Percentage change in total revenues (*%ΔRevenue*)	26,374	9.26	−1.69	7.10	16.87	27.87
Percentage change in program spending (*%ΔProgram Ratio*)	26,374	−1.16	−8.14	0.21	8.65	29.13
Firm-Specific Regression Outputs						
Compensation:						
Intercept (β_0)		4.9560	2.4771	5.3488	7.9415	7.4726
Change in revenues (β_1)		0.1963	−0.0679	0.0654	0.3955	0.7289
Change in program ratio (β_2)		0.4135	0.0000	0.0992	0.4328	0.9123
Adjusted R^2		0.0978	−0.1625	0.0254	0.2969	0.3250
Donations:						
Intercept (β_0)		2.2919	−0.8775	6.9014	9.7973	20.2525
Change in revenues (β_1)		0.8002	−0.0046	0.0884	1.1558	2.3809
Change in program ratio (β_2)		0.8655	0.0000	0.0759	0.5724	2.7266
Adjusted R^2		0.2268	−0.1355	0.0803	0.6087	0.4260

This table presents the firm-specific estimates of the sensitivity of changes in compensation (*%ΔComp*) and donations (*%ΔDons*) to changes in the program ratio (*%ΔProgram Ratio*). The firm-specific sensitivity estimates (β_2) measure a nonprofit's relative incentive to manage its program ratio. The sample is from the NCCS database and includes the years 1985 to 2001. We require that each firm report non-zero values for all variables, have eight years of valid data, and not report any government or feeder donations. The compensation model is from Baber et al. (2002) while the donations model is a straightforward extension.

Variable Definitions:

%ΔComp = the one-year percentage change in officer's compensation, line 25a IRS 990;

%ΔDons = the one-year percentage change in public donations, line 1a IRS 990;

%ΔRevenue = the one-year percentage change in total revenues, line 12 IRS 990; and

%ΔProgram Ratio = the one-year percentage change in the ratio of program expenses to total revenues, multiplied by revenue at time t and deflated by program expenses at time $t-1$ (see Baber et al. 2002). Program expenses are those spent on charitable (as opposed to administrative or fundraising)

year of SFAS Nos. 116 and 117, both of which changed the format of nonprofit financial statements from a fund accounting orientation to a commercial accounting orientation. SFAS No. 117 had several provisions, including requiring specific financial statements and specific forms of those statements, which could affect the reported amounts of expenses. SFAS No. 117 requires that material amounts of expenses (program, administrative, and fundraising) be stated separately in the financial statements or notes, a requirement that could decrease the probability that a nonprofit would misreport its expenses.[15] The primary purpose of SFAS No. 116 was to clarify the definition of a "contribution" (versus other types of revenues resulting from "exchange or agency" transactions such as membership dues or sponsorships) and thus, because of matching, also necessarily addresses if and when to record the related "fundraising expense." Although SFAS Nos. 116 and 117 did not have a direct effect on how the IRS 990 is prepared, to the extent that audited financial statements are used as a basis for preparing the IRS 990, the passage of the new accounting rules could have had an effect on IRS 990 reporting. Indeed, Jones and Roberts (2006) find that most nonprofits that undergo an audit on their financial statements prepare their IRS 990 in accordance with SFAS Nos. 116 and 117 when those rules do not disagree with IRS 990 preparation rules.

We include a second indicator variable equal to 1 if the year is greater than or equal to 1998, the effective date of the American Institute of Certified Public Accountant's Statement of Position 98-2 (SOP 98-2), which changed the way nonprofits report the expenses associated with combined fundraising and educational activities. Prior to SOP 98-2, many nonprofits classified all costs related to these joint campaigns as programs, which had the effect of inflating the program ratio and deflating the fundraising ratio (Jones and Roberts 2006). SOP 98-2 requires nonprofits to partition these costs between programs and fundraising. Roberts (2005) and Jones and Roberts (2006) examine the effects of SOP 98-2 on nonprofit financial reporting and in general find that it had a positive effect on nonprofit financial reporting quality.

In addition to incentives and accounting rule changes, inherent expertise could play a role in the accuracy of a nonprofit's financial information. Although it is difficult to directly measure the construct of expertise, we include the use of an outside accountant (as reported on line 31a of the IRS 990) as a reasonable proxy.[16] We recognize that an alternative interpretation of this variable is that the use of an outside accounting firm could indicate that the nonprofit received a financial statement audit, and the resulting audited financial statements could provide the basis for a more complete and correct IRS 990. We therefore interpret this variable with caution. As discussed later, we also partition our models across size and age to calibrate the effects of potential expertise that is associated with larger or older nonprofits, as age and size could be measures of inherent expertise or general level of care.

We also consider the role of financial need in the reporting decision. In theory, a nonprofit with more pressing financial need would be more inclined to manipulate their reported financial ratios so as to garner additional donations. However, there are several difficulties with empirically measuring the concept of financial distress. First, we recognize that it is difficult to capture timing effects. If an organization in financial distress manipulates its financial information and then receives additional donations in the subsequent period, then in those subsequent periods there would not necessarily be a relationship between financial need and reporting choices as the financial need would ostensibly be overcome. In an attempt to mitigate this timing effect we use a lagged value of financial need, although this is an incomplete method as the number of lags is unknown (we do not know how long ago the organization first manipulated its financial reports as a response to financial distress). Second, it is not clear how one would measure the construct of financial need in the nonprofit setting. To this end we rely on a working paper by Keating et al. (2005) who find that the ratio of net assets to total assets is a predictor of nonprofit financial distress measured as either insolvency risk (negative net equity) or financial disruption risk (a 25 or more percent decline in net assets).[17] Because of the theoretical problems discussed above we estimate our model separately with and without the financial need metric to ensure that our primary results are not affected by the inclusion of the financial need variable.

---

[15] For example, SFAS No. 117 states that donations must not be netted against outside fundraising costs. If a nonprofit obtains all of its donations from outside fundraising firms, prior to SFAS No. 117 they could have correctly reported *net* donations (i.e., net of fundraising costs paid to the outside fundraising firm). However, after SFAS No. 117 the nonprofit must report *gross* donations and any associated fundraising costs. This rule change could have an impact on the propensity to report fundraising expenses on IRS Form 990.

[16] We chose to use a dichotomous variable rather than a continuous one (i.e., payment to outside accountant) because the relevant partition is whether or not the organization hires an outside accountant at all, and not on how much it pays that accountant. Certainly larger firms will pay their outside accountants more, but this is not necessarily an economically meaningful relationship. Nonetheless, we tested the robustness of our results to the inclusion of the continuous variable (payment to accountant) with no change in the tenor of the results.

Additional control variables include *Size* (total assets from line 59b of the IRS 990 divided by one million) and a trend variable equal to 1 in 1985, 2 in 1986 and so on. The trend variable is intended to control for increases in nonprofit financial reporting quality across time, ensuring that the accounting change variables are independent of general increases in nonprofit reporting quality over time. The results are not sensitive to the inclusion of additional control variables such as total donations, total revenues, and the square of size.[18]

Empirical Model

Our empirical model takes the general form:

$$\begin{aligned} \mathit{Misreported\ Expenses}_{it} = {} & \alpha + \beta_1 \mathit{Compensation\ Sensitivity}_{it} \\ & + \beta_2 \mathit{Donations\ Sensitivity}_{it} + \beta_3 \mathit{Outside\ Accountant}_{it} \\ & + \beta_4 \mathit{SFAS116/117}_{it} + \beta_5 \mathit{SOP\ 98\text{-}2}_{it} + \beta_6 \mathit{Size}_{it} \\ & + \beta_7 \mathit{Trend}_{it} + \beta_8 \mathit{Financial\ Need}_{it-1} + \varepsilon_{it}. \end{aligned} \tag{3}$$

To ensure that our statistical inferences are not influenced by the effects of using panel data, which has the tendency to deflate the standard errors of the regression estimates and increase the chance of finding significant relationships, we estimate our models using two methods that control for the effects of non-independence across time. First, we cluster the standard errors at the firm level. Second, we use annual Fama-Macbeth regressions, where statistical significance is based on the distribution of the annual coefficient estimates. As noted by Petersen (2004), each of these methods has its relative strengths and weaknesses.

---

[17] Keating et al. (2005) find that several other ratios are also indicators of nonprofit financial distress. We used each of their measures and found mixed results. While some of their financial distress measures are associated with the probability of reporting zero fundraising expenses (i.e., net assets/total assets and net assets/total revenues), others are not (i.e., net income/total revenue and the annual change in net income).

[18] In Table 3 we find that compensation and donations are related to changes in program spending that arise from changes in revenues and changes in program ratios. While managers may seek to increase their revenue so as to garner higher compensation and donations, the link between this incentive and expense misreporting is not clear, hence we exclude these incentive variables (the β_1s from Equations (1) and (2)) from our incentive tests. However, the results of estimating Equation (3) are robust to including these additional variables.

The Fama-MacBeth approach is more standard in the accounting and finance literatures, but it does not allow us to include the variables for accounting rule changes, whereas the corrected standard error approach does. We present our clustered standard error results as they allow us to use a more complete model, and discuss our Fama-MacBeth results.

Our methodology involves two stages; first, we estimate our sensitivity measures by regressing compensation or donations on the program ratio, and second, we test for incentive effects by regressing expense misreporting (as a zero fundraising expense indicator) on our sensitivity measures. Because program expenses appear in the first stage and are related to fundraising expenses (total expenses are the sum of program, fundraising, and administrative expenses), while fundraising expenses appear in the second stage (as the dependent variable), it might appear that our sensitivity variables are mechanically correlated with fundraising introducing endogeneity bias into our results. For example, if a nonprofit systematically reports zero fundraising expenses (our dependent variable in equation 3), then it is likely to have systematically higher program expenses (a variable used in Models 1 and 2 to derive the sensitivity variables). However, because we estimate our sensitivity metrics using changes at the firm level in the first stage (Equations (1) and (2), our variables are free of any mechanical relationship with the dependent variable in the second stage (i.e., Equation (3)). Even if a nonprofit systematically overstates its program expenses by reporting zero fundraising expenses, any percentage changes in its reported program expenses, holding fundraising expenses a constant at zero, will be free of mechanical bias.[19]

Table 4 contains the descriptive statistics for the incentive analysis variables. The dependent variable, *Misreported Expenses,* is an indicator variable equal to 1 if the organization reported zero fundraising expenses, and 0 otherwise. We use this measure of expense misreporting based on the results earlier in the analysis, which provided evidence that many of the firms that report zero fundraising expenses are under-reporting those expenses. Certainly this metric is not without error, as undoubtedly many firms that report some fundraising expenses on their 990 under-report those expenses and it is possible that some firms that report zero fundraising on their 990 truly incur zero fundraising expenses. In both of these cases we misclassify our dependent variable causing measurement error. However, unless this error in the dependent variable is correlated with our independent variables our results will be unbiased. As shown in Table 4, approximately 50 percent of our sample reports zero fundraising expenses, despite earning public donations. The compensation and donations sensitivity variables are discussed earlier in this section and reported in Table 3. The statistics of the incentive variables are slightly different across Tables 3 and 4 because the sample in Table 4 is limited to those observations with usable data for all of the regression variables used in Equation (3). Approximately 82 percent of our observations hire an outside accountant, while the average nonprofit is 36 years old.

Next we discuss the results of estimating Equation (3). Because the underlying distribution of some of the variables appears to be skewed, we also estimated our models using ranks with no change in the statistical inferences of the results.

Regression Results

Table 5 contains the results of the analysis for the full sample as well as industry partitions (i.e., charitable, educational, and medical). Model 1 of Table 5 reports results for

---

[19] To ensure that our results are not affected by observations that report zero fundraising in some years and positive amounts in others, we re-estimated our results on the sample of observations that report positive or zero fundraising in all sample years, with no change in the inferences of the study.

TABLE 4
Descriptive Statistics Analysis of Characteristics Associated with Understated Fundraising Expenses

Variable	Mean	First Quartile	Median	Third Quartile	Standard Deviation
Misreported Expenses	0.50	0.00	1.00	1.00	0.50
Compensation Sensitivity	0.42	0.00	0.09	0.39	0.97
Donations Sensitivity	0.94	0.00	0.09	0.55	2.94
Outside Accountant	0.82	1.00	1.00	1.00	0.38
SFAS 116/117	0.54	0.00	1.00	1.00	0.50
SOP 98-2	0.30	0.00	0.00	1.00	0.46
Size	0.95	0.17	0.42	0.93	2.43
Trend	9.88	7.00	10.00	13.00	3.83
Financial Need	0.72	0.54	0.82	0.96	0.28
Age	36.01	20.98	36.88	50.91	17.57

The sample includes the 5,780 observations in the NCCS database years 1985 to 2001 that have non-missing values for all variables.

Variable Definitions:

Misreported Expenses = equal to 1 if the organization reported zero fundraising expenses, and 0 otherwise;
Outside Accountant = equal to 1 if any fees were paid to outside accountants (a non-zero amount on line 31a IRS 990), and 0 otherwise;
SFAS 116/117 = an indicator is equal to 1 post-1994, and 0 otherwise;
SOP98-2 = an indicator is equal to 1 post-1997, and 0 otherwise;
Size = total assets (in $100 millions);
Trend = a time trend variable equal to 1 in 1985, 2 in 1986, and so on; and
Financial Need = lagged ratio of net assets to total assets.

Compensation and *Donations Sensitivity* variables are described and derived in Table 3.

the full sample (using clustered standard errors) and shows that the probability of reporting zero fundraising expenses is increasing in the relative sensitivities of the program ratio to both officer's compensation and donations. This finding provides support for our agency-based predictions and is consistent with the notion that nonprofit managers intentionally manipulate their publicly available financial information so as to attract higher compensation and donations. We also find that the use of an outside accountant reduces the probability that a nonprofit will report zero fundraising expenses, consistent with the use of an outside paid accountant increasing the reliability and usefulness of nonprofit financial reports. Finally, we find that although SOP 98-2 reduces the probability of a nonprofit reporting zero fundraising expenses, the enactment of SFAS No. 116/117 appears to have no effect on misreported fundraising expenses. This finding is consistent with the results of Jones and Roberts (2006), who document that SOP 98-2 had an effect on how nonprofits report the results of combined fundraising and educational campaigns.

Models 2, 3, and 4 in Table 5 show little variation in these general results across the three primary nonprofit industries (i.e., charitable, educational, and medical nonprofits).[20] The only exception to the full sample is that the use of outside paid accountants has no effect on expense misreporting by charitable nonprofits (although the sign of the coefficient is in the correct direction). Model 5 includes the financial need variable for the full sample

---

[20] Educational nonprofits include primary and secondary nongovernmental schools, medical nonprofits include primary care hospitals, and charitable nonprofits include all other types of charities.

TABLE 5
Regression Analysis of Characteristics Associated with Understated Fundraising Expenses (t-statistics Clustered by Firm)

$$\begin{aligned} \textit{Misreported Expenses}_{it} = {} & \alpha + \beta_1 \textit{Compensation Sensitivity}_{it} + \beta_2 \textit{Donations Sensitivity}_{it} \\ & + \beta_3 \textit{Outside Accountant}_{it} + \beta_4 \textit{SFAS No. 116/117}_{it} \\ & + \beta_5 \textit{SOP 98-2}_{it} + \beta_6 \textit{Size}_{it} + \beta_7 \textit{Trend}_{it} \\ & + \beta_8 \textit{Financial Need}_{it-1} + \varepsilon_{it}. \end{aligned}$$

	Column				
	1	2	3	4	5
Sample Composition	**Full Sample**	**Charities**	**Educational**	**Medical**	**Full Sample**
Intercept	−0.1887	−0.0965	−1.6028	1.5012	0.5581
	(0.4058)	(0.7926)	(0.0003)	(0.0095)	(0.0695)
Compensation Sensitivity	0.6772	0.4834	1.3999	1.1952	0.6088
	(<.0001)	(0.0128)	(0.0001)	(0.0045)	(<.0001)
Donations Sensitivity	0.8470	0.8786	1.2236	0.9260	0.8061
	(<.0001)	(<.0001)	(0.0010)	(0.0015)	(<.0001)
Outside Accountant	−0.3607	−0.0857	−1.0303	−0.8998	−0.2994
	(0.0464)	(0.7669)	(0.0047)	(0.0460)	(0.0974)
SFAS116/117	0.0291	−0.0048	−0.1116	0.3012	0.0026
	(0.7100)	(0.9710)	(0.4421)	(0.1371)	(0.9742)
SOP98-2	−0.2293	−0.3123	−0.6678	−0.5252	−0.2090
	(0.0043)	(0.0151)	(0.0042)	(0.0342)	(0.0061)
Size	−0.0744	−0.1476	−0.0778	−0.3082	−0.0754
	(0.0121)	(0.1541)	(0.0364)	(0.0035)	(0.0350)
Trend	−0.0024	−0.0360	−0.0379	0.0305	0.0073
	(0.8954)	(0.2014)	(0.3303)	(0.5074)	(0.7055)
Financial Need					−1.1289
					(<.0001)
Observations	5,771	2,226	1,677	1,719	5,655

The sample includes the 5,780 observations in the NCCS database years 1985 to 2001 that have non-missing values for all variables.

Two-tailed *p*-values are in parentheses.

Variable Definitions:

Misreported Expenses = equal to 1 if the organization reported zero fundraising expenses, and 0 otherwise;
Outside Accountant = equal to 1 if any fees were paid to outside accountants, and 0 otherwise;
SFAS116/117 = an indicator is equal to 1 post-1994, and 0 otherwise;
SOP 98-2 = an indicator is equal to 1 post-1997, and 0 otherwise;
Size = total assets (in $100 millions);
Trend = a time trend variable equal to 1 in 1985, 2 in 1986, and so on; and
Financial Need = the lagged ratio of net assets over total assets. The educational sample includes private universities and secondary schools, the medical sample includes primary care hospitals, and the charity sample includes all other types of nonprofits.

Compensation and *Donations sensitivity* variables are described and derived in Table 3.

and shows that, as suspected, nonprofits with more pressing financial needs (as evidenced by a lower ratio of net assets to total assets) are more likely to report zero fundraising expenses (i.e., a negative coefficient suggests that a lower ratio of net assets to total assets is associated with a higher probability of reporting zero fundraising expenses).

We estimated similar tests using the Fama-MacBeth method where coefficients are averaged across annual regressions and the corresponding t-statistics are based on the distribution of the annual coefficients. The inferences of the results (untabulated) using the Fama-MacBeth method are identical to those in Table 5. The consistency and robustness of our results provides comfort that they are not driven by intemporal correlations within firms.

We also estimated the models using the Fama-MacBeth approach for various size and age groupings. The groupings are the lower quartile, middle half, and upper quartile groupings of size and age. We group on these variables because it is possible that the effects of the independent variables on financial statement management could vary across size and age in ways that are not fully captured by simply including an age or size regressor. For example, it is possible that smaller nonprofits are inherently less interested in the quality of their financial reports, or they lack the necessary expertise, either hired from the outside or employed on the inside, to prepare correct IRS 990s. Our results (untabulated) show that our results hold across size and age partitions, suggesting that neither size nor age drives the overall results. Results are similar using the clustered standard error method.

To summarize the results of this portion of our analysis, our findings are consistent with agency theory and suggest that nonprofits that have compensation and resource (i.e., donation revenues) incentives to manage their financial reports respond to those incentives. Our empirical analyses are consistent even when we use a variety of statistical models, sample partitions, and various definitions of the incentive variables. The finding that nonprofit managers respond to incentives to manage their publicly available financial information has important policy implications as it suggests that nonprofits are providing donors and regulators with managed financial information that could possibly lead to incorrect donations allocations and governing decisions.

V. CONCLUSIONS

IRS Form 990 is the principle source of financial information for nonprofits, and the relative allocation of expenses among the categories on the IRS 990 is of primary interest to donors and other stakeholders. A growing body of anecdotal evidence notes the fact that over 50 percent of all nonprofits that receive donations also report zero fundraising expenses. It is widely suspected that this behavior is inappropriate, and perhaps even intentional, as a nonprofit can increase the amount of program expenses it reports to the public on its IRS 990 by shifting expenses from fundraising to programs. Our analysis provides empirical evidence supporting the notion that nonprofits inappropriately and perhaps even intentionally misstate their relative proportion of expenses in these three categories.

We provide evidence suggesting that many nonprofits reporting zero fundraising expenses do, in fact, undertake fundraising activities. Further analysis shows that expense misreporting is associated with managerial incentives to garner larger amounts of managerial pay and donations.

Our findings have important implications for public policy to the extent that donors and other stakeholders allocate their resources among nonprofits or make regulating decisions based on their reported financial results. Our results also have important implications for nonprofit researchers as they suggest that organizations that report zero fundraising expenses should be carefully considered. Depending on the type of analysis and question being asked it might be prudent for a researcher to exclude organizations from their analysis that report zero fundraising expenses.

REFERENCES

Arrow, K. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53: 941–973.

Baber, W. R., A. A. Roberts, and G. Visvanathan. 2001. Charitable organizations' strategies and program-spending ratios. *Accounting Horizons* 15: 329–343.

———, P. L. Daniel, and A. A. Roberts. 2002. Compensation to managers of charitable organizations: An empirical study of the role of accounting measures of program activities. *The Accounting Review* 77: 679–694.

Callen, J. 1994. Money donations, volunteering, and organizational efficiency. *The Journal of Productivity Analysis* 5: 215–228.

Fama, E., and M. Jensen. 1983. Agency problems and residual claims. *Journal of Law and Economics* 26: 301–325.

Froelich, K., T. Knoepfle, and T. Pollak. 2000. Financial measures in nonprofit organization research: Comparing IRS 990 return and audited financial statement data. *Nonprofit and Voluntary Sector Quarterly* 29: 232–254.

Greenlee, J., and K. Brown. 1999. The impact of accounting information on contributions to charitable organizations. *Research in Accounting Regulation* 13: 111–125.

Guidry, F., A. J. Leone, and S. Rock. 1999. Earnings-based bonus plans and earnings management by business-unit managers. *Journal of Accounting and Economics* 26: 113–142.

Hager, M. 2003. Current practices in allocation of fund-raising expenditures. *New Directions for Philanthropic Fund-Raising* 41 (December): 39–52.

Hansmann, H. 1996. *The Ownership of Enterprise*. Cambridge, MA: Belknap Press.

Harvey, J., and K. McCrohan. 1988. Fundraising costs—Societal implications for philanthropies and their supporters. *Business and Society* 27: 15–22.

Healy, P. 1985. The effects of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7: 85–107.

Holthausen, R., D. Larcker, and R. Sloan. 1995. Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics* 19: 29–74.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics* 4: 469–506.

Joint Committee on Taxation (JCT). 2000. Study of present-law taxpayer confidentiality and disclosure provisions as required by section 3802 of the Internal Revenue Service Restructuring and Reform Act of 1998. Volume I. January. Washington, DC: Government Printing Office.

Jones, C., and A. A. Roberts. 2006. Management of financial information in charitable organizations: The case of joint cost allocations. *The Accounting Review* 81 (1): 159–178.

Keating, E., M. Fischer, T. Gordon, and J. Greenlee. 2005. Assessing financial vulnerability in the nonprofit sector. Working paper, Harvard University.

Khumawala, S., and T. Gordon. 1997. Bridging the credibility of GAAP: Individual donors and the new accounting standards for nonprofit organizations. *Accounting Horizons* 11: 45–68.

Okten, C. and B. A. Weisbrod. 2000. Determinants of donations in private nonprofit markets. *Journal of Public Economics* 75: 255–272.

Petersen, M. 2004. Estimating standard errors in finance panel data sets: Comparing approaches. Working paper, Northwestern University.

Posnett, J., and T. Sandler. 1989. Demand for charity donations in private non-profit markets. *Journal of Public Economics* 40: 187–200.

Roberts, A. A. 2005. The implications of joint cost standards for charity reporting. *Accounting Horizons* 19: 11–23.

Sansing, R., and R. Yetman. 2006. Governing private foundations using the tax law. *The Journal of Accounting and Economics* (forthcoming).

Tinkelman, D. 1999. Factors affecting the relation between donations to not-for-profit organizations and an efficiency ratio. *Research in Governmental and Nonprofit Accounting* 10: 135–61.

United States General Accounting Office (GAO). 2002. *Report to the Chairman and Ranking Minority Member, Committee on Finance, U.S. Senate.* GAO–02–526. Washington, D.C.: Government Printing Office.

Weisbrod, B., and N. Dominguez. 1986. Demand for collective goods in private nonprofit markets: Can fund-raising expenditures help overcome free-rider behavior? *Journal of Public Economics* 30: 83–95.

Yetman, M. H., and R. J. Yetman. 2004. The effects of governance on the financial reporting quality of nonprofit organizations. Working paper, University of California, Davis.

Yetman, R. J., and M. H. Yetman. 2003. The effect of nonprofits' taxable activities on the supply of private donations. *National Tax Journal* 56: 243–58.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 421–441

Income Smoothing over the Business Cycle: Changes in Banks' Coordinated Management of Provisions for Loan Losses and Loan Charge-Offs from the Pre-1990 Bust to the 1990s Boom

Chi-Chun Liu
National Taiwan University

Stephen G. Ryan
New York University

ABSTRACT: Prior research shows that during the pre-1990 bust financially weak banks managed income upward by delaying provisions for losses on heterogeneous loans. In contrast, we predict and find that during the 1990s boom profitable banks managed income downward by accelerating provisions for losses on homogeneous loans. Profitable banks obscured their income smoothing by accelerating charge-offs of homogeneous loans and by recording more gross charge-offs to offset recoveries of previously charged-off loans. Over the three years subsequent to the acceleration of charge-offs, they had higher and more persistent income before provisions for loan losses than other banks, consistent with income smoothing over a prolonged horizon.

Keywords: *income smoothing; business cycle; banks; provisions for loan losses; loan charge-offs.*

Data availability: *All data are available from public sources.*

I. INTRODUCTION

We investigate how commercial banks' management of provisions for loan losses and loan charge-offs changed from their period of poor health and ultimately crisis (bust) from 1974–1990 to their period of robust health (boom) throughout the 1990s. A considerable body of prior research finds that during the pre-1990 bust banks delayed provisioning for loan losses, with this effect being stronger for financially weaker banks holding more heterogeneous loans. In contrast, we predict and find that during the

---

The authors thank Bruce Bettinghaus for helping us access the regulatory data for commercial bank holding companies. The authors thank two anonymous referees, Anwer Ahmed, Dan Gode, Jim Ohlson, Jim Wahlen, and accounting seminar participants at Georgetown University, National Taiwan University, New York University, and The University of Texas at Austin for useful comments. Professor Liu gratefully acknowledges the financial support of the National Science Council, Taiwan, Republic of China (NSC 91-2416-H-002-028).

Editor's note: The paper was accepted by Dan S. Dhaliwal, Senior Editor.

Submitted May 2003
Accepted October 2005

1990s boom banks accelerated provisions and made this less apparent by also accelerating charge-offs, with these effects being stronger for more profitable banks holding more homogeneous loans.

As in Liu and Ryan (1995), we exploit differences in how provisions and charge-offs are determined for heterogeneous and homogeneous loans under Generally Accepted Accounting Principles and related accounting practices. For heterogeneous (e.g., commercial) loans, provisions and charge-offs typically are determined with judgment on a loan-by-loan basis. For homogeneous (e.g., consumer) loans, provisions are determined using statistical methods at the portfolio level, while charge-offs are determined using number-of-days-past-due rules. Liu and Ryan (1995) predict and find that banks had more ability to *delay* provisions for *heterogeneous* than homogeneous loans during the pre-1990 bust. In contrast, we predict and find that banks had more ability to *accelerate* charge-offs for *homogeneous* than heterogeneous loans during the 1990s boom.

We focus on banks' acceleration of charge-offs of homogeneous loans to obscure their income smoothing, a previously undocumented discretionary behavior, not on the strategic reasons for that behavior. Numerous studies argue convincingly that banks have incentives to smooth their income, and in particular to store income in booms such as the 1990s to cushion income in subsequent busts. Consistent with these incentives, most of the numerous empirical studies investigating this question—on samples drawn wholly or largely from the pre-1990 bust—find that banks do smooth their income using provisions for loan losses. Building on this prior research, we posit that banks prefer smoother income.

We also posit that, in smoothing their income, banks prefer that their allowances for loan losses do not fluctuate too much, in order to avoid scrutiny by bank regulators, the SEC, auditors, and market participants. The primary way that banks can reconcile these preferences is by exercising discretion over loan charge-offs, for example, by accelerating charge-offs when the allowance for loan losses would otherwise be too high. Bank regulators and analysts facilitate this discretionary behavior by using the ratio of the allowance for loans losses to net loan charge-offs to assess the adequacy of banks' allowances for loan losses.

We test and find support for four specific hypotheses about banks' provisioning and charge-off behavior during the 1990s boom. First, we predict and find that more profitable banks holding more homogeneous loans exhibit a more positive association between the provision for loan losses and income before the provision for loan losses, consistent with greater income smoothing by these banks. Second, we predict and find that more profitable banks charging off more homogeneous loans recover those charge-offs at a higher rate, consistent with greater acceleration of charge-offs by these banks. Third, we predict and find that more profitable banks recovering more charge-offs of homogeneous loans in a given year record more gross loan charge-offs in that year, consistent with these banks making new charge-offs to keep their allowances for loan losses from ballooning. Fourth, we predict and find that banks charging off more homogeneous loans had higher and more persistent income before the provision for loan losses over the next three years than did other banks, consistent with these banks smoothing income over a prolonged horizon.

Collectively, our results strongly suggest that profitable banks managed their income downward and accelerated charge-offs of homogeneous loans to obscure this discretionary behavior over the prolonged 1990s boom.[1] Our paper is the first to provide evidence of banks' management of loan charge-offs for the purposes of income smoothing; prior research typically treats charge-offs as a nondiscretionary control variable, perhaps because charge-offs do not directly affect income. It is also the first to provide evidence of banks' management of provisions and charge-offs for homogeneous rather than heterogeneous loans in boom times. This evidence adds to our understanding of the range and determinants of banks' discretionary behavior.

Our paper also contributes to the broader area of research on income smoothing. As discussed by Dechow and Skinner (2000), a debate remains in this literature—especially between academics and practitioners—as to whether, when, and how income smoothing occurs. In our view, a primary reason for the lack of resolution of this debate is that income smoothing is a more contextual and less purely symmetrical behavior than is captured by the usual research designs. We provide an example of such behavior, showing that banks manage income downward in boom times differently from how they manage income upward in bust times.

The remainder of this paper is organized as follows. Section II develops our hypotheses based on the rules and practices governing the accounting for loan losses and prior research. Section III describes the sample and reports descriptive statistics. Section IV reports the empirical results. Section V concludes.

II. HYPOTHESIS DEVELOPMENT

Accounting for and Disclosures of Loan Losses

Under FAS No. 5 (FASB 1975), *Accounting for Contingencies*, when credit losses on a loan or portfolio of homogeneous loans are probable and can be reasonably estimated, an expense called the provision for loan losses and a contra-asset (to loans outstanding) called the allowance for loan losses should be recorded. When a specific loan is deemed uncollectible, it should be charged off, which involves reducing both loans outstanding and the allowance for loan losses by the uncollectible amount. Charge-offs do not directly affect total assets, net loans outstanding, owners' equity, or net income. Charged-off loans may turn out to be collectible, in which case the charge-offs are reversed and called recoveries. Net loan charge-offs equal gross loan charge-offs less recoveries.

Banks are required to disclose nonperforming assets, which includes the entire book value of loans more than 90 days past due, restructured troubled debt, and repossessed real estate. Nonperforming assets is a relatively nondiscretionary and timely source of information about loan default, although it does not reflect either the probability of default or expected losses given default.

Banks' Discretion over Provisions for Loan Losses

The Effect of the 1990s Boom

This study was motivated in part by our belief that banks saw their income increase in the early 1990s and immediately began overstating provisions for loan losses, initially to undo past under-reserving, but fairly quickly (no later than 1994) over-reserving. Consistent with this belief, from no later than October 1994 throughout the remainder of the 1990s boom, various governmental and regulatory bodies repeatedly expressed concern that banks were over-reserved for loan losses and thereby smoothing their income across the business cycle. Most notably, in an October 1994 report, the General Accounting Office states that a sample of "institutions maintained significant amounts of unsupported reserves ... not clearly linked to likely losses" (U.S. GAO 1994). In November 1998, the SEC required SunTrust Banks—which reported a high level of gross charge-offs that it recovered at a rate close to 40 percent from 1993–1997—to reduce its allowance for loan losses by $100 million. The public nature of this requirement was intended as a signal to all banks either to justify or reduce their allowances.[2] In January 1999, the SEC sent a form letter to selected bank holding companies suggesting specific enhanced disclosures regarding their allowance and provision for loan losses. In July 2001, the SEC staff issued SAB No. 102, *Selected Loan Loss Allowance Methodology and Documentation Issues*, which requires "a systematic methodology to be employed each period in determining the allowance for loan losses to be reported," emphasizing that this methodology should be "consistently applied" through time. The major bank regulators contemporaneously revised their regulatory guidance to be consistent with SAB No. 102.

---

[1] While not directly related to this paper, Moyer (1990), Beatty et al. (1995), and Collins et al. (1995) make hypotheses about and generally find evidence of the management of regulatory capital using net loan charge-offs. Loan charge-offs reduce regulatory capital to the extent that the allowance for loan losses is included in regulatory capital. The allowance for loan losses was included in primary capital up to 1990 and subsequently is included with limits in Tier 2 capital.

Prior Research

Prior research on banks' management of provisions for loan losses falls into three distinct literatures. The first argues that banks have incentives to decrease income in good times such as the 1990s boom to allow increased income in subsequent bad times. Joyce (1996) argues that banks manage provisions for compensation purposes. Galai et al. (2003) argue that banks create "hidden reserves" that they release as needed either to raise income to meet analyst forecasts or capital requirements. Kanagaretnam et al. (2003) hypothesize that banks that smooth income downward are more likely to obtain financing externally.

The second literature examines empirically whether banks manage provisions for loan losses to smooth income. Most, though not all, of these studies find evidence consistent with income smoothing.[3] Kanagaretnam et al. (2003) find that downward income smoothing is stronger for banks that are more likely to raise funds externally. Beatty et al. (2002) find that publicly traded banks are more likely than privately held banks to manage earnings upward to avoid earnings decreases.

The third literature examines the market pricing of the allowance or provision for loans losses.[4] All of these studies are motivated at least in part by the idea that banks exercise discretion over provisions for loan losses, with several advancing the hypothesis that banks do so to signal their quality.[5] A number of these studies examine the market's reaction to the March–April 1987 announcements of provisions for less-developed-country debt for which banks had clearly exercised discretion to delay making those provisions, given that much of this debt was known to be troubled since the early 1980s.[6]

---

[2] The "SunTrust Banks—After the Restatement" case in Ryan (2002, Chapter 5) examines this situation in detail.

[3] Greenawalt and Sinkey (1988), Ma (1988), Collins et al. (1995), Robb (1998), Beatty et al. (2002), Kanagaretnam et al. (2003), and Gray (2004) provide evidence that banks smooth income using provisions for loan losses, while Scheiner (1981), Beatty et al. (1995), and Ahmed et al. (1999) do not.

[4] See Beaver et al. (1989), Wahlen (1994), Liu and Ryan (1995), Beaver and Engel (1996), Liu et al. (1997), Ahmed et al. (1999), and the papers on less-developed-country debt cited in footnote 6.

[5] Beaver et al. (1989), Musumeci and Sinkey (1990), Grammatikos and Saunders (1990), Wahlen (1994), and Liu et al. (1997) advance the signaling hypothesis.

[6] See Grammatikos and Saunders (1990), Musumeci and Sinkey (1990), Elliott et al. (1991), and Griffin and Wallach (1991). Poland and other Eastern bloc countries had problems repaying debt in the early 1980s and Mexico and Brazil announced moratoria repaying interest and principal in late 1982. Boehmer and Megginson (1990) report that the average market price for less-developed-country debt in January 1986 ranged from 27.5 cents on the dollar for Peru to 83 cents on the dollar for Columbia.

Effects of Loan Portfolio Composition

Liu and Ryan (1995) argue that, under Generally Accepted Accounting Principles and related accounting practices, banks have differential ability to exercise discretion over provisions for different types of loans. For homogeneous (e.g., consumer) loans, provisions are determined primarily on a statistical basis at the portfolio level. For heterogeneous (e.g., commercial) loans, provisions are determined primarily on a judgmental basis at the individual-loan level. During bust periods when banks have incentives to delay provisions for loan losses, the statistical approach should yield timelier and less discretionary provisions for homogeneous loans than the judgmental approach yields for heterogeneous loans. Consistent with this argument, for a sample of banks during the pre-1990 bust, Liu and Ryan (1995) find that provisions for loan losses were highly untimely compared to nonperforming assets for banks that held above-median heterogeneous loans, but relatively timely for banks that held above-median homogeneous loans.

In contrast, we expect discretionary acceleration of both provisions and charge-offs for homogeneous loans during the 1990s boom. Homogeneous loans typically are charged off based on number-of-days-past-due rules, while heterogeneous loans typically are charged off based on judgment on a loan-by-loan basis. While the use of number-of-days-past-due rules may suggest that less discretion is applied to charge-offs of homogeneous than heterogeneous loans, this turns out not to be true when banks desire to accelerate charge-offs in boom times, for three reasons.

First, these rules interact with economic conditions, yielding charge-offs that are recovered with much higher probability in good than bad times. While this could be viewed as a mechanical rather than discretionary aspect of these rules, in principle banks could and should adjust the number of days past due they use in these rules to remove this effect. In practice, however, they do not appear to do so, perhaps reflecting McNichols and Wilson's (1988) observation that as "receivables turnover decline (for example, in banking), mechanical rules are more difficult to specify, leaving greater room for discretion over write-offs."

Second, and more importantly, at least until bank regulators changed the guidance for charge-offs of consumer loans effective June 1999, as discussed below, banks employed discretion in determining and applying their policies regarding the number of days past due to charge off homogeneous loans, leading to time-series and cross-sectional variation in those policies.[7] To curb diversity across banks in charge-off practices for homogeneous loans, in February 1999 the Federal Financial Institutions Examination Council (FFIEC, a cooperative group of the main bank regulators) revised guidance for charge-offs of consumer loans that had been in place since 1980, stating that "[p]revious policy guidance had been interpreted and applied inconsistently," "[c]harge-off practices ranged from 120 to

---

[7] Until pressured by the SEC to do so in late 1998 (via the SunTrust restatement) and early 1999 (via the sample letter sent to bank holding companies [SEC 1999]), banks rarely disclosed their charge-off policies in their public filings. Hence, during our sample period, it is generally not possible to determine when and how banks changed their charge-off policies. We determined that a fair number of such changes did occur by conducting a web search using "charge-off policy" and "change." For example, Standard & Poor's Creditwire explained the change in charge-off rates for an index of credit card-backed securities from June to July 1997 as being significantly affected by First USA's decision to charge off credit card accounts at 180 days past due rather than 210 days past due it had used previously.

240 days" past due, and "a number of institutions were not following existing policy guidance for charging off open-end accounts based on past due status."[8]

Third, charge-offs of heterogeneous loans typically are based on loan-by-loan evaluations by loan officers, whose incentives generally work against acceleration of charge-offs and whose decisions are harder to coordinate. Office of the Comptroller of the Currency et al. (1993) states that "an effective credit grading framework generally places primary reliance on loan officers to identify emerging loan problems." However, in a 2001 speech, Federal Reserve Board Chairman Alan Greenspan stated "regrettably, incentives for loan officers ... traditionally have rewarded loan growth, market share, and the profits that derive from booking interest income with, in retrospect, inadequate provisions for possible default." Udell (1989) and Berger and Udell (2002) discuss and provide evidence about loan officers' informational advantages in monitoring loans and their incentive to hide loan default on the loans they originated. They argue that loan officers devote insufficient time to monitoring existing loans, because their compensation is based disproportionately on short-term revenues. Berger and Udell (2003) emphasize that this time allocation issue and other incentive problems are more likely to arise in booms, such as the 1990s, when the opportunities for loan growth are highest and when realized loan default has been low for extended periods.

Hypotheses

Building on the prior research discussed above, we posit that banks prefer smoother income, all else being equal. During the 1990s boom, smoothing usually involved banks reducing income, as evidenced by the actions of the governmental and regulatory bodies described above. We also posit that, in smoothing their income, banks prefer that their allowances for loan losses do not fluctuate too much, in order to avoid scrutiny by bank regulators, the SEC, auditors, analysts, and other market participants. We hypothesize that banks satisfy both these preferences by exercising discretion over loan charge-offs. During the 1990s boom, this usually involved banks accelerating loan charge-offs to reduce the allowance for loan losses.

While banks' loan charge-offs are observable, this discretionary behavior likely is opaque to outsiders for the following two reasons. First, charge-offs can be volatile at the bank level even in good times, and discretionary charge-offs are not directly observable. Hence, the exercise of discretion can appear to be attributable to chance, at least for a while. Second—and more importantly given the prolonged period over which we hypothesize that banks smoothed income—bank regulators and analysts facilitate this discretionary behavior by using the ratio of the allowance for loans losses to net loan charge-offs to assess the adequacy of banks' allowances for loan losses.[9] Increasing net charge-offs lowers this ratio by both reducing its numerator and increasing its denominator, so a bank that that increases its charge-offs appears to have a less excessive allowance for loan losses.

We test four hypotheses about how more profitable banks smoothed their income downward during the 1990s boom by accelerating provisions for and charge-offs of homogeneous loans for which charge-offs are determined using number-of-days-past-due rules and so can be more easily accelerated. First, consistent with most prior empirical research, which finds that banks smooth their income, we hypothesize that banks' provisions for loan losses are positively associated with income before the provision for loan losses. Extending this prior research, we hypothesize that this association is stronger for more profitable banks that hold more homogeneous loans, consistent with greater income smoothing by these banks.

---

[8] See FFIEC (1999) February 10, 1999. In this guidance, the FFIEC requires credit card and other open-end consumer loans to be charged off no later than 180 days past due and closed-end consumer (e.g., auto) loans to be charged off no later than 120 days past due as of June 1999. These new rules do not eliminate differences in charge-off practices across banks, however, because they allow banks to adopt a "policy more conservative than the one detailed." For example, SunTrust Banks continues to charge off closed-end consumer loans at 90 (not 120) days past due, despite consistently recovering these charge-offs at about a 40 percent rate.

[9] See Federal Reserve Board (2005, Section 2072.1, 13), Office of the Comptroller of the Currency (1998, 38), and Ryan (2002, Chapter 5).

Second, we hypothesize that more profitable banks that charged off more homogeneous loans had higher rates of recoveries of charge-offs, consistent with these banks accelerating charge-offs of these loans. As discussed above, recovery rates are used by bank regulators and analysts to assess the aggressiveness of banks' charge-off policies.[10]

Third, we hypothesize that more profitable banks that recovered more charge-offs of homogeneous loans in a given year recorded more new gross loan charge-offs in that year, consistent with these banks reducing their allowances for loan losses in order to obscure their income smoothing across the prolonged 1990s boom.

Fourth, we predict and find that profitable banks that charged off more homogeneous loans have higher and more persistent income before the provision for loan losses over the next three years, consistent with these banks smoothing income over a prolonged horizon.

III. SAMPLE AND DESCRIPTIVE STATISTICS

The sample covers the period from 1991–2000, which includes the entire 1990s boom. The number of commercial bank holding companies in the sample rises from 122 in 1991 to 304 in 2000, peaking at 329 in 1999. Since the sample does not include the bust period up to 1990, we contrast our results to those of the prior research summarized in Section II to make inferences about changes over time in banks' management of provisions for loan losses and loan charge-offs. While this is not optimal, we view it as preferable to attempting to construct a sample of banks that includes the bust period for the following reasons. First, the bank regulatory data discussed below becomes available only in 1986, and its definitions of variables change at the beginning of our sample period in significant ways. In particular, the data by loan type discussed below became available in its current form either at the end of 1990 or in the first quarter of 1991. Second, Bank Compustat does not include non-surviving banks, and so yields a highly selected sample in the bust period. Third, there are nontrivial problems matching banks over a longer period given the volume of mergers and acquisitions in the banking industry.

We obtain the following data by type of loan from commercial bank holding companies' regulatory Y-9C reports available on the Federal Reserve Bank of Chicago's website: loans outstanding, gross and net loan charge-offs, recoveries, and nonperforming assets. As in Liu and Ryan (1995), homogeneous loans include consumer loans, 1-4 family residential mortgages, loans to financial institutions, and acceptances of other banks. Heterogeneous loans include commercial and industrial loans, direct lease financing, all other real estate loans, agriculture loans, and foreign loans.[11] We also obtain the allowance and provision for loan losses, net income, total assets, and owners' equity from these reports.

---

[10] See Federal Reserve Board (2005, Section 2072.1, 13), Office of the Comptroller of the Currency (1998, 17), and Ryan (2002, Chapter 5).

[11] Liu and Ryan (1995) use more aggregated data on loans outstanding by type from Bank Compustat to calculate homogeneous and heterogeneous loans. The main difference between our classification scheme and that of Liu and Ryan (1995) is our more descriptive classification of 1-4 family residential mortgages as homogeneous, whereas Liu and Ryan are constrained to classify all real estate loans in a single category, which they choose to be heterogeneous. This classification difference has minimal effect on the results in this paper.

We obtain Tier 1 risk-based capital ratios from 1993–2000 from Bank Compustat, because these ratios are only available beginning in 1997 in the Y-9C reports. Since Tier 1 risk-based capital ratios are not available prior to 1993 on Bank Compustat, in those years we proxy for it using equity divided by assets.

Table 1 reports descriptive statistics for the variables in the regression analyses reported in Section IV as well as additional descriptive variables. Panel A reports the means and quartiles of these variables for the entire sample period.

Panel B of Table 1 reports medians of the variables for the odd years in the sample period. The trends in the medians from 1991 to 1995 indicate that by no later than 1994 banks were experiencing low levels of loan losses and were far better reserved for loan losses than during the pre-1990 bust. Specifically, the downward trends for the provision for loan losses (*PLL*) and net loan charge-offs (*NLCO*) both deflated by beginning loans outstanding, indicate that loan losses dropped by about 65 percent from 1991 to 1995. The upward trends for the ratios of the allowance for loan losses (*ALL*) to nonperforming assets (*NPA*) and to *NLCO* indicate that loss reserve adequacy improved by about 150 percent over this period. The ratio of recoveries to gross loan charge-offs, *REC%*, referred to as the recovery rate, is a measure of the aggressiveness of banks' charge-off policies used by bank regulators and analysts. The upward trend for *REC%* indicates that the recovery rate doubled from 1991 to 1995, rising to a level of 32.5 percent in 1995 that is difficult to reconcile with charged-off loans being deemed uncollectible.

Reflecting banks' improving health, their return on assets, *ROA*, rose almost 40 percent from 0.96 percent in 1991 to 1.33 percent in 1995 and stayed at about that level throughout the remainder of the 1990s. The upward trend for the Tier 1 risk-based capital ratio, *CAP*, reported in Panel B is correct but overstated, because the ratio of book equity to book assets proxies for this variable in 1991. Bassett and Zakrajsek (2001) report that the average Tier 1 risk-based capital ratio for all U.S. chartered commercial banks rose approximately 30 percent, from .081 in 1991 to 0.106 in 1994, and Panel B indicates that our sample banks stayed well capitalized through 2000. These statistics illustrate a highly profitable and well-capitalized banking industry by no later than 1994, conditions that persist through 2000.

Panel C of Table 1 reports medians of $REC\%_t$ for groups of observations more and less likely to accelerate loan charge-offs to obscure their income smoothing. Using the prior year's values of variables, we group observations into above- and below-median ROA_{t-1} groups, into quintiles based on total gross loan charge-offs divided by loans outstanding, $GLCO_{t-1}$, and into quintiles based on each of gross loan charge-offs of homogeneous and heterogeneous loans divided by loans outstanding, $GLCO(HOM)_{t-1}$ and $GLCO(HET)_{t-1}$, respectively. Consistent with our hypothesis that more profitable banks with higher charge-offs of homogeneous loans are more likely to be accelerating loan charge-offs, for the above-median ROA_{t-1} group, $REC\%_t$ rises sharply and monotonically from 21.6 percent in the lowest $GLCO(HOM)_{t-1}$ quintile to 34.6 percent in the highest $GLCO(HOM)_{t-1}$ quintile. In contrast, $REC\%_t$ has no obvious trend across the $GLCO(HOM)_{t-1}$ quintiles for the below-median ROA_{t-1} group. Also in contrast, $REC\%_t$ generally decreases across the $GLCO(HET)_{t-1}$ quintiles for the above-median ROA_{t-1} group, and it has no obvious trend across these quintiles for the below-median ROA_{t-1} group.

Panel D of Table 1 reports Pearson correlations of these variables, which indicate that many of our explanatory variables—most notably, the provision for loan losses and net loan charge-offs—are highly correlated. While these correlations do not pose problems in

TABLE 1
Descriptive Statistics[a,b,c,d]

Panel A: Means and Quartiles of Variables in Pooled Sample

Variables	Mean	Min	25%	Med	75%	Max	#Obs
X	.0223	−.0117	.0177	.0216	.1260	.0822	2148
ROA	.0127	−.0197	.0103	.0128	.0155	.0392	1886
PLL	.0048	−.0041	.0019	.0033	.0053	.0476	2148
NLCO	.0038	−.0026	.0011	.0023	.0044	.0417	2148
REC	.0013	.0000	.0005	.0009	.0026	.0127	2159
REC%	.3575	.0029	.1614	.2558	.4033	4.5223	2137
ΔNPA	−.0004	−.0612	−.0020	.0000	.0021	.0413	2148
HOM%	.4731	.0361	.3633	.4775	.5959	.9708	2145
ALL/NPA	2.9521	.1921	1.0458	1.9008	3.3003	50.2022	2105
ALL/NLCO	8.0720	−62.2102	3.0375	5.4541	10.4358	77.7020	1751
CAP	.1156	.0414	.0850	.1099	.1370	.3071	1736

Panel B: Medians of Variables in Odd Years of Sample Period

Variables	1991	1993	1995	1997	1999
X	.0219	.0232	.0219	.0225	.0204
ROA	.0096	.0129	.0133	.0134	.0124
PLL	.0075	.0040	.0029	.0032	.0029
NLCO	.0058	.0028	.0022	.0021	.0018
REC	.0012	.0012	.0010	.0008	.0006
REC%	.1576	.2912	.3253	.2648	.2474
ΔNPA	.0014	−.0023	−.0003	.0002	.0003
HOM%	.4763	.5085	.5064	.4854	.4663
ALL/NPA	.7615	1.2770	1.9847	2.2635	2.4508
ALL/NLCO	2.5246	5.3861	6.2509	5.8624	6.1907
CAP	.0730	.1139	.1122	.1100	.1051
max #obs	122	141	197	252	329

Panel C: Median Recovery Rate ($REC\%_t$) by Profitability and Charge-Off Groups

		Quintiles				
ROA_{t-1}	**Quintiles Based on Variables**	**1 (low)**	**2**	**3**	**4**	**5 (high)**
Above median	$GLCO_{t-1}$	.2198	.2101	.2370	.2646	.2658
	$GLCO(HOM)_{t-1}$	.2158	.2244	.2579	.3007	.3457
	$GLCO(HET)_{t-1}$	.3529	.2663	.2271	.2371	.2146
Below median	$GLCO_{t-1}$	.2894	.2960	.2832	.2707	.2607
	$GLCO(HOM)_{t-1}$	.2647	.3563	.2467	.3000	.2560
	$GLCO(HET)_{t-1}$	.2831	.2717	.2389	.2850	.2563

(*continued on next page*)

TABLE 1 (Continued)

Panel D: Pearson Correlations of Variables in Pooled Sample

Variables	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
X (1)	.90	.32	.26	.17	−.00	−.05	−.17	.15	−.06	−.06
ROA (2)		−.11	−.06	.10	.06	−.05	−.10	.22	.05	−.05
PLL (3)			.86	.33	−.27	−.02	−.10	−.03	−.23	−.02
NLCO (4)				.41	−.33	−.18	−.04	−.06	−.26	−.04
REC (5)					.36	−.23	−.08	−.04	−.24	−.04
REC% (6)						−.02	−.04	.10	−.04	.04
ΔNPA (7)							.02	−.00	.01	.02
HOM% (8)								.00	.13	−.05
ALL/NPA (9)									.18	.06
ALL/NLCO (10)										.73
CAP (11)										

---

[a] All variables are collected from commercial bank holding companies' regulatory Y-9C filings except for Tier 1 risk-based capital, which is collected from Bank Compustat. The sample covers the period 1991–2000.

[b] The descriptive statistics reflect all observations of any variable included in any subsequent regression analysis. Observations in the outside 0.5 percent of either tail of the distribution of any variable in a given analysis or for which total assets or loans outstanding on the regulatory Y-9C reports differs from the value on Bank Compustat by more than 10 percent are deleted.

[c] Variable definitions:

X = net income before the provision for loan losses divided by beginning of year total loans;
ROA = net income divided by beginning of year total assets;
PLL = provision for loan losses divided by beginning of year total loans;
GLCO = gross loan charge-offs divided by beginning of year total loans;
NLCO = net loan charge-offs divided by beginning of year total loans;
REC = recoveries divided by beginning of year total loans;
REC% = recoveries divided by gross loan charge-offs;
ΔNPA = change in nonperforming assets divided by beginning of year total loans;
HOM% = percentage of total loans outstanding that are homogeneous loans (consumer loans, 1–4 family residential mortgages, loans to financial institutions, or acceptances of other banks);
ALL/NPA = allowance for loan losses divided by nonperforming assets;
ALL/NLCO = allowance for loan losses divided by net loan charge-offs;
CAP = Tier 1 risk-based capital ratio; and
max #obs = the maximum number of observations of any variable in a given year (which typically exceeds the minimum number of observations of any variable in that year by 10–20 percent).

[d] A variable name followed by *HOM* or *HET* in parentheses denotes the value of that variable for homogeneous or heterogeneous loans, respectively. For example, *GLCO(HOM)* denotes gross loan charge-offs of homogeneous loans.

our pooled estimations, they preclude reliable year-by-year estimations of the models, especially in the early years in our sample period that have about one-third the observations of later years.

IV. MODELS AND RESULTS

Since our single-industry sample exhibits cross-sectional dependence, we estimate each regression model with fixed time effects. To mitigate heteroscedasticity, we deflate unscaled variables by beginning loans outstanding and delete observations of each variable in the outside 0.5 percent of each tail of its distribution each year. Our regression results are not sensitive to the treatment of outliers as long as extreme observations are deleted or pulled in; for example, deletion of influential outliers identified based on Belsley et al. (1980) diagnostics has minimal effect on our results.

Income Smoothing Using Provisions for Loan Losses

In this section, we test the hypothesis that during the 1990s boom, banks' provisions for loan losses are positively associated with their income before the provision for loan losses, consistent with banks managing provisions to smooth their income, and that this association is stronger for more profitable banks that hold more homogeneous loans. We conduct these tests using the following variant of a model developed by Ahmed et al. (1999):[12]

$$PLL_t = a + bI_{highroa_t} + cHOM\%_t + dX_t + f(X_t{*}I_{highroa_t}) + g(X_t{*}HOM\%_t) + hCAP_t + j\Delta NPA_t + e_t. \quad (1)$$

$I_{highroa,t}$ is an indicator variable that takes a value of 1 if a bank has an above-median return on assets in year t, and 0 otherwise. $HOM\%_t$ denotes the percentage of homogeneous loans in the bank's loan portfolio. X_t denotes income before the provision for loan losses.[13]

The hypothesis stated above has three main implications for the coefficients in Equation (1). First, the coefficient on X_t should be positive, consistent with banks smoothing their income using provisions for loan losses during the 1990s boom. Second, the coefficient on $X_t{*}I_{highroa,t}$ should be positive, consistent with more profitable banks more aggressively smoothing income (downward) during this period. Third, the coefficient on $X_t{*}HOM\%_t$ should be positive, consistent with banks that hold more homogeneous loans more aggressively smoothing income during this period.

While our hypothesis pertains to the coefficients on the variables in which $I_{highroa,t}$ and $HOM\%_t$ are interacted with X_t, we also include $I_{highroa,t}$ and $HOM\%_t$ separately in Equation (1) to ensure that the coefficients on the interaction terms do not reflect any direct effects of profitability and loan portfolio composition on provisions for loan losses. Ahmed et al. (1999) include CAP_t to control for banks recording provisions to raise regulatory capital and ΔNPA_t to control for economic loan default during the period, and they find a negative coefficient on CAP_t and a positive coefficient on ΔNPA_t.

PLL_t, X_t, and ΔNPA_t are deflated by loans outstanding at the beginning of year t. The other variables in Equation (1) are naturally deflated or scale-free.

Table 2 reports the fixed time effects estimation of Equation (1). As predicted, the coefficient on X_t is significantly positive (t = 1.8), consistent with banks smoothing their income using provisions for loan losses during the 1990s boom. As predicted, the coefficient on $X_t{*}I_{highroa,t}$ is significantly positive (t = 3.5), consistent with more profitable banks more aggressively smoothing income using provisions for loan losses than less profitable banks during the 1990s boom. The coefficient on $I_{highroa,t}$ is significantly positive (t = 2.9), also

---

[12] Equation (1) differs from the model in Ahmed et al. (1999) in the following ways. First, we do not include a difficult-to-collect control variable, the change in per capita dollar liabilities of failed businesses weighted by the geographic distribution of loans, which they found to be not significant. Second, we do not include the implied standard deviation of return on assets, which they use to proxy for the change in the risk of the loan portfolio, because this variable cannot be calculated for almost 20 percent of the sample due to insufficient returns data. We did conduct sensitivity analysis adding this variable for the subsample for which it is available, however, and the conclusions from our hypothesis tests are not altered. Third, we do not include certain interaction terms involving X and CAP that relate to their time-period-specific hypotheses. Fourth, we include $I_{highroa,t}$ by itself and interacted with X to test hypotheses regarding bank profitability. Fifth, we include $HOM\%$ by itself and interacted with X to test hypotheses regarding homogeneous loans.

[13] For simplicity, we use income before the pretax provision for loan losses throughout the paper. We tested the sensitivity of our results to using earnings before 0.6 times the provision for loan losses, and found no significant differences in the conclusions from our hypothesis tests (the coefficients on the provision for loan losses are algebraically affected, of course).

TABLE 2

Regression of the Provision for Loan Losses on Earnings before the Provision for Loan Losses with Profitability and Loan Portfolio Composition Interactions[a,b,c,d]

$$PLL_t = a + bI_{highroa_t} + cHOM\%_t + dX_t + f(X_t{*}I_{highroa_t}) + g(X_t{*}HOM\%_t) + hCAP_t + j\Delta NPA_t + e_t. \qquad (1)$$

	Hypotheses	**Coefficients**
$I_{highroa,t}$		.008* (2.9)
$HOM\%$		−.011 (−1.6)
X_t	+	.124* (1.8)
$X_t{*}I_{highroa,t}$	+	.718* (3.5)
$X_t{*}HOM\%$	+	.573* (2.2)
CAP_t		.015 (1.2)
ΔNPA_t		−.039 (−1.5)
R^2		.27
# obs		1739

* White's heteroscedasticity-adjusted t-statistic (shown in parentheses) is significant at the 5 percent level or better.

[a] The sample is described in the notes to Table 1.

[b] All variables are defined in the notes to Table 1 except for $I_{highroa,t}$, which takes a value of 1 if an observation has above median return on assets in year *t*, and 0 otherwise.

[c] The equation is estimated with fixed time effects, which are not tabulated.

[d] The R^2 does not include the explanatory power of the fixed effects.

---

possibly reflecting banks' income smoothing. As predicted, the coefficient on $X_t{*}HOM\%_t$ is significantly positive (t = 2.2), consistent with the hypothesis that this income smoothing is stronger for banks with more homogeneous loans.

In contrast to Ahmed et al.'s (1999) finding of significant coefficients on CAP_t and ΔNPA_t, we find none of the coefficients on the control variables CAP_t, ΔNPA_t, and $HOM\%_t$ are significant. This lack of significance suggests that banks determined their provisions for loan losses primarily based on their profitability during the 1990s boom, perhaps because they were well capitalized and experienced low levels of loan default during that period.

In summary, in this section we find that the provision for loan losses is positively correlated with income before the provision for loan losses, consistent with banks smoothing their income using provisions for loan losses during the 1990s boom. We find this income smoothing behavior is stronger for more profitable banks holding more homogeneous loans.

Acceleration of Loan Charge-Offs

In this section, we test the hypothesis that more profitable banks accelerated charge-offs of homogeneous loans in order to keep their overstated allowances for loan losses from ballooning during the 1990s boom. We conduct these tests using the following model:

$$\begin{aligned} REC_t = a &+ bHOM\%_{t-1} + cI_{highroa_{t-1}} + dGLCO(HET)_{t-1} \\ &+ f[GLCO(HET)_{t-1}{*}I_{highroa_{t-1}}] + gGLCO(HOM)_{t-1} \\ &+ h[GLCO(HOM)_{t-1}{*}I_{highroa_{t-1}}] + jCAP_{t-1} + \sum_{s=0}^{3} k_s \Delta NPA_{t-s} + e_t \end{aligned} \qquad (2)$$

We infer that banks have charged off loans more quickly when their gross loan charge-offs have a more positive association with one-year-ahead recoveries, denoted REC_t.[14] We distinguish gross loan charge-offs for heterogeneous and homogeneous loans, denoted $GLC(HET)_{t-1}$ and $GLCO(HOM)_{t-1}$, respectively, because we expect banks to have more ability to accelerate charge-offs of homogeneous loans than heterogeneous loans. Because we expect that more profitable banks accelerated charge-offs more than less profitable banks during the 1990s boom, we interact these gross loan charge-off variables with $I_{highroa,t}$.

The coefficients on the gross loan charge-off variables in Equation (2) reflect the rates of recovery in year t of incremental charge-offs in year t–1. Since some level of recoveries is to be expected even in the absence of discretion, the coefficients on both $GLCO(HET)_{t-1}$ and $GLCO(HOM)_{t-1}$ should be positive. The hypothesis stated above has two main implications for these coefficients. First, the coefficient on $GLCO(HOM)_{t-1}{*}I_{highroa,t-1}$ should be positive, consistent with more profitable banks accelerating charge-offs for homogeneous loans more than less profitable banks. Second, the sum of the coefficients on $GLCO(HOM)_{t-1}$ and $GLCO(HOM)_{t-1}{*}I_{highroa,t-1}$ should be larger than the sum of the coefficients on $GLCO(HET)_{t-1}$ and $GLCO(HET)_{t-1}\ {*}I_{highroa,t-1}$, i.e., $g + h > d + f$, consistent with more profitable banks accelerating charge-offs more for homogeneous loans than for heterogeneous loans.

We include $HOM\%_{t-1}$ and $I_{highroa,t-1}$ separately in Equation (2) to ensure that the coefficients on the interaction variables do not reflect any direct effects of loan portfolio composition and profitability on recoveries. We include CAP_t because Moyer (1990), Beatty et al. (1995), and Collins et al. (1995) predict and generally find that loan charge-offs are used to manage regulatory capital.[15] We include current and three lagged changes in nonperforming assets, ΔNPA_{t-s}, $0 \leq s \leq 3$, to control for banks' economic exposure to loan losses, because nonperforming assets is the best available measure of those losses.

REC_t, $GLCO(HET)_{t-1}$, $GLCO(HOM)_{t-1}$, and ΔNPA_{t-s} are deflated by loans outstanding at the beginning of year $t-3$. The other variables in Equation (2) are naturally deflated or scale-free.

Table 3 reports the fixed time effects estimation of Equation (2). As expected even in the absence of discretion, the coefficients on $GLCO(HET)_{t-1}$ and $GLCO(HOM)_{t-1}$ are significantly positive (t = 7.1 and 11.4, respectively). These coefficients imply that gross charge-offs for heterogeneous loans are recovered 11.9 percent of the time, while gross charge-offs for homogeneous loans are recovered at the much higher rate of 23.8 percent.

As predicted, the coefficient on $GLCO(HOM)_{t-1}{*}\ I_{highroa,t-1}$ is significantly positive (t = 2.2), consistent with more profitable banks accelerating charge-offs of homogeneous

---

[14] We are able to demonstrate excessive loan charge-offs by healthy banks during the 1990s in various ways other than the association between gross loan charge-offs and future recoveries (e.g., by a rising ratio of net loan charge-offs to nonperforming assets). We chose recoveries as our dependent variable because of bank regulators' and analysts' use of recoveries to assess the aggressiveness of banks' loan charge-off policies, as discussed in Section II.

[15] We do not expect loan charge-offs to be delayed to increase regulatory capital after the first couple of years in our sample period, however, because virtually all banks are well capitalized by no later than 1994.

TABLE 3
Regression of Recovery Rate on Heterogeneous and Homogeneous Loan Charge-Offs with Profitability Interactions[a,b,c,d]

$$\begin{aligned} REC_t = {} & a + bHOM\%_{t-1} + cI_{highroa_{t-1}} + dGLCO(HET)_{t-1} \\ & + f[GLCO(HET)_{t-1} {*} I_{highroa_{t-1}}] + gGLCO(HOM)_{t-1} \\ & + h[GLCO(HOM)_{t-1} {*} I_{highroa_{t-1}}] + jCAP_{t-1} + \sum_{s=0}^{3} k_s \Delta NPA_{t-s} + e_t. \end{aligned} \qquad (2)$$

	Hypotheses	Coefficients
$HOM\%$		−.014 (−.4)
$I_{highroa,t-1}$		−.007 (−.8)
$GLCO(HET)_{t-1}$		.119* (7.1)
$GLCO(HET)_{t-1} {*} I_{highroa,t-1}$		.046 (1.4)
$GLCO(HOM)_{t-1}$		.238* (11.4)
$GLCO(HOM)_{t-1} {*} I_{highroa,t-1}$	+	.083* (2.2)
CAP_{t-1}		.010* (3.4)
ΔNPA_t		−1.642* (−2.8)
ΔNPA_{t-1}		−2.610* (−4.1)
ΔNPA_{t-2}		−2.017 (−3.1)
ΔNPA_{t-3}		−1.479 (−2.1)
g+h−d−f	+	.156 (3.6)
R^2		.47
# obs		1709

---

* White's heteroscedasticity-adjusted t-statistic (shown in parentheses) is significant at the 5 percent level or better.

[a] The sample is described in the notes to Table 1.

[b] All variables are defined in the notes to Tables 1 and 2.

[c] The equation is estimated with fixed time effects, which are not tabulated.

[d] The R^2 does not include the explanatory power of the fixed effects.

loans more than other banks. This result implies that above-median-profitability banks recovered gross loans charge-offs of homogeneous loans at an 8.3 percent higher rate than below-median-profitability banks. Also as predicted, the difference between the sum of the coefficients on $GLCO(HOM)_{t-1}$ and $GLCO(HOM)_{t-1}{*}\ I_{highroa,t-1}$ and the sum of the coefficients on $GLCO(HET)_{t-1}$ and $GLCO(HET)_{t-1}{*}\ I_{highroa,t-1}$ is significantly positive (t = 3.6), consistent with more profitable banks accelerating charge-offs of homogeneous loans more than heterogeneous loans.

The coefficient on the control variable CAP_{t-1} is significantly positive (t = 3.4), implying better capitalized banks recorded charge-offs more quickly than did poorly capitalized banks.[16] The coefficients on ΔNPA_{t-s}, $0 \leq s \leq 3$, are all significantly negative, consistent with these variables capturing economic loan default. The coefficients on $HOM\%_{t-1}$ and $I_{highroa,t-1}$ are not statistically significant.

In summary, the descriptive statistics in Table 1, Panel B discussed above show that banks experienced sharply increasing recovery rates during the early 1990s, consistent with banks charging off loans too quickly. Consistent with the descriptive analysis of recovery rates in Table 1, Panel C, in this section we find that, superimposed on this overall trend, recovery rates were higher for more profitable banks holding more homogeneous loans. These results are consistent with loan charge-offs being substantially determined by more profitable banks' discretionary behavior during the 1990s boom.

Sustaining the Income Smoothing

An alternative explanation for the high recovery rates for profitable banks charging off homogeneous loans is that these banks were consistently surprised by the persistently favorable economic conditions during the extended 1990s boom, more so than other banks. In this section, we attempt to rule out this alternative explanation and demonstrate how these banks obscured their income smoothing over such an extended period. Specifically, we test the hypothesis that during the 1990s boom profitable banks experiencing recoveries of homogeneous loans recorded new gross loan charge-offs in order to reduce their allowances for loan losses.

To test this hypothesis, we estimate the following model:

$$\begin{aligned} GLCO_t = {} & a + bHOM\%_t + cI_{highroa_t} + dREC(HET)_t + f[REC(HET)_t{*}I_{highroa_t}] \\ & + gREC_t(HOM)_t + h[REC(HOM)_t{*}I_{highroa_t}] \\ & + jCAP_t + kPLL_t + lALL_{t-1} + m\Delta NPA_t + e_t \qquad (3) \end{aligned}$$

The model captures the association of gross loan charge-offs with contemporaneous recoveries of charge-offs of both heterogeneous and homogeneous loans, $REC(HET)_t$ and $REC(HOM)_t$, respectively, distinguishing above- and below-median-profitability banks.

The coefficients on the recovery variables in Equation (3) reflect the rates of new gross loan charge-offs per incremental dollar of recoveries. The hypothesis stated above has two main implications. First, the coefficient on $REC(HOM)_t{*}I_{highroa,t}$ should be positive, consistent with more profitable banks recording more new charge-offs when they experience recoveries of homogeneous loans than do less profitable banks. Second, the sum of the coefficients on $REC(HOM)_t$ and $REC(HOM)_t{*}I_{highroa,t}$ should be larger than the sum of the coefficients on $REC(HET)_t$ and $REC(HET)_t{*}I_{highroa,t}$, i.e., $g + h > d + f$, consistent with more profitable banks recording more new gross loan charge-offs when they experience recoveries of homogeneous rather than heterogeneous loans.

We include $HOM\%_t$ in Equation (3) to control for loan portfolio composition. We include $I_{highroa,t}$ separately to ensure that the coefficients on the interaction variables do not capture any direct effect of profitability. We include CAP_t to control for the previously discussed finding of prior research that banks use loan charge-offs to manage regulatory capital. We include PLL_t and ΔNPA_t to control for new information about loan default during the year. We include ALL_{t-1} in the model to control for prior estimates of loan default.

---

[16] This result could be attributable to more poorly capitalized banks delaying charge-offs to increase regulatory capital, as the research discussed above generally finds happened during the bust period prior to 1990. Given that virtually all the banks in our sample are well capitalized, however, it is more likely that the better capitalized banks accelerated charge-offs.

$GLCO_t$, $REC(HET)_t$, $REC(HOM)_t$, PLL_t, ALL_{t-1}, and ΔNPA_t are deflated by loans outstanding at the beginning of year t. The other variables in Equation (3) are naturally deflated or scale-free.

Table 4 reports the fixed time effects estimation of Equation (3). As predicted, the coefficient on $REC(HOM)_t{*}I_{highroa,t}$ is significantly positive (t = 1.9), consistent with more profitable banks recording more new charge-offs when they experience recoveries of homogeneous loans than less profitable banks. Also as predicted, the difference between the

TABLE 4

Regression of Gross Loan Charge-Offs on Recoveries of Heterogeneous and Homogeneous Loans with Profitability Interactions[a,b,c,d]

$$\begin{aligned} GLCO_t = {} & a + bHOM\%_t + c\,I_{highroa_t} + dREC(HET)_t + f[REC(HET)_t * I_{highroa_t}] \\ & + gREC_t(HOM)_t + h[REC(HOM)_t * I_{highroa_t}] + \\ & + jCAP_t + kPLL_t + lALL_{t-1} + m\Delta NPA_t + e_t \end{aligned} \tag{3}$$

	Hypotheses	**Coefficients**
$HOM\%$		−.027 (−1.5)
$I_{highroa,t}$		.035* (2.5)
$REC(HET)_t$		.547 (.8)
$REC(HET)_t{*}I_{highroa,t}$		1.022 (1.0)
$REC(HOM)_t$		2.980* (5.5)
$REC(HOM)_t{*}I_{highroa,t}$	+	1.468* (1.9)
CAP_t		−.099 (−.8)
PLL_t		.478* (4.0)
ALL_{t-1}		.059 (1.6)
ΔNPA_t		−.022 (−.3)
g+h−d−f	+	2.879* (1.8)
R^2		.97
#obs		1729

---

* White's heteroscedasticity-adjusted t-statistic (shown in parentheses) is significant at the 5 percent level or better.

[a] The sample is described in the notes to Table 1.

[b] All variables are defined in the notes to Tables 1 and 2.

[c] The equation is estimated with fixed time effects, which are not tabulated.

[d] The R^2 does not include the explanatory power of the fixed effects.

sum of the coefficients on $REC(HOM)_t$ and $REC(HOM)_t{*}I_{highroa,t}$ and the sum of the coefficients on $REC(HET)_t$ and $REC(HET)_t{*}I_{highroa,t}$ is significantly positive (t = 1.8), consistent with more profitable banks recording more new gross loan charge-offs when they recover homogeneous loans than when they recover heterogeneous loans.

The coefficient on $I_{highroa,t}$ is significantly positive (t = 2.5), likely reflecting some of the hypothesized discretionary behavior. The coefficient on PLL_t is significantly positive (t = 4.0), reflecting the well-known high positive correlation between charge-offs and provisions for loan losses. The coefficients on $HOM\%_t$, CAP_t, ALL_{t-1}, and ΔNPA_t are all not statistically significant.

In summary, in this section we find that during the 1990s boom profitable banks that experienced more recoveries of charge-offs of homogeneous loans recorded more new gross loan charge-offs in the same year. These results are consistent with the hypothesis that banks recorded those loan charge-offs to reduce their allowances for loan losses and thereby obscure their downward income smoothing during that period.

The Level and Persistence of Net Income

Due to the prolonged nature of the 1990s boom, profitable banks smoothing their income downward had to do so over a fairly (and perhaps unexpectedly) long horizon. In this section, we test the hypothesis that profitable banks that charged off more homogeneous loans during this period had higher and more persistent income before the provision for loan losses over the next three years, consistent with them smoothing income over a prolonged horizon.

We test this hypothesis using the following model:

$$\begin{aligned} X_{t+s} = {} & a + bHOM\%_t + c\,NLCO(HET)_t + dNLCO(HOM)_t + fX_t \\ & + g(X_t * HOM\%_t) + h[X_t * NLCO(HET)_t] + j[X_t * NLCO(HOM)_t] \\ & + kCAP_t + lPLL_t + m\Delta NPA_t + e_{t+s},\ 1 \le s \le 3. \end{aligned} \tag{4}$$

This model regresses income before the provision for loan losses for each of the next three years, X_{t+s}, $1 \le s \le 3$, on net loan charge-offs for both homogeneous and heterogeneous loans, denoted $NLCO(HET)_t$ and $NLCO(HOM)_t$, respectively, both separately and interacted with current income before the provision for loan losses, X_t.

Based on the hypothesis stated above, we expect positive coefficients on $NLCO(HOM)_t$, indicating that higher net charge-offs of homogeneous loans are associated with higher future income before the provision for loan losses. We also expect positive coefficients on $X_t{*}NLCO(HOM)_t$, indicating that higher net charge-offs for homogeneous loans are associated with more persistent income before the provision for loan losses.

We include X_t in Equation (4) to capture the average level of income persistence for the sample. We include $HOM\%_t$ and $X_t{*}HOM\%_t$ to control for the effect of loan portfolio composition on the level and persistence of income. We include CAP_t to capture the previously discussed finding of prior research that banks use loan charge-offs to manage regulatory capital. We include PLL_t and ΔNPA_t to control for new information about loan default during the year.

The variables X_{t+s}, $0 \le s \le 3$, $NLCO(HET)_t$, $NLCO(HOM)_t$, PLL_t, and ΔNPA_t are deflated by loans outstanding at the beginning of year *t*. The other variables in Equation (4) are naturally deflated.

Table 5 reports the fixed time effects estimation of Equation (4). As predicted, the coefficient on $NLCO(HOM)_t$ is significantly positive for the two- and three-year horizons

TABLE 5

Regression of Future Earnings before the Provision for Loan Losses on Current Net Loan Charge-Offs of Heterogeneous and Homogeneous Loans and Earnings before the Provision for Loan Losses with Interactions between Charge-Offs and Earnings[a,b,c,d]

$$\begin{aligned} X_{t+s} = {} & a + bHOM\%_t + cNLCO(HET)_t + dNLCO(HOM)_t + fX_t + g(X_t * HOM\%_t) \\ & + h[X_t * NLCO(HET)_t] + j[X_t * NLCO(HOM)_t] + kCAP_t + lPLL_t \\ & + m\Delta NPA_t + e_{t+s},\ 1 \leqslant s \leqslant 3. \end{aligned} \tag{4}$$

	Hypotheses	**Coefficients**		
		X_{t+1}	X_{t+2}	X_{t+3}
$HOM\%_t$		−.001	−.001	−.004
		(−1.0)	(−.3)	(−1.3)
$NLCO(HET)_t$		−.186	.227	.211
		(−1.1)	(1.1)	(.8)
$NLCO(HOM)_t$	+	.142	.711*	.663*
		(.9)	(3.0)	(2.4)
X_t		.952*	1.087*	1.240*
		(29.2)	(21.5)	(15.9)
$X_t*HOM\%_t$		−.491*	−.607*	−.591*
		(−3.0)	(−3.2)	(−2.0)
$X_t*NLCO(HET)_t$		−6.167	−31.892	−56.546*
		(−1.2)	(−1.5)	(−4.0)
$X_t*NLCO(HOM)_t$	+	19.718*	45.234*	82.764*
		(2.7)	(2.7)	(3.5)
CAP_t		−.003	−.009	−.018
		(−1.0)	(−1.3)	(−1.5)
PLL_t		.050	−.600*	−.763*
		(.3)	(−3.7)	(−3.0)
ΔNPA_t		−.023	−.015	−.126
		(−.6)	(−.4)	(−1.7)
R^2		.69	.53	.41
# obs		1680	1363	1115

* White's heteroscedasticity-adjusted t-statistic (shown in parentheses) is significant at the 5 percent level or better.

[a] The sample is described in the notes to Table 1.

[b] All variables are defined in the notes to Tables 1 and 2.

[c] The equation is estimated with fixed time effects, which are not tabulated.

[d] The R^2s do not include the explanatory power of the fixed effects.

(t = 3.0 and 2.4, respectively), consistent with banks recording more charge-offs of homogeneous loans having higher future income before provisions for loan losses than other banks, though this coefficient is positive but not significant for the one-year horizon. In contrast, the coefficient on $NLCO(HET)_t$ is not statistically significant. Also as predicted, the coefficient on $X_t*NLCO(HOM)_t$ is significantly positive for each of the one-year to three-year horizons (t = 2.7, 2.7, and 3.5, respectively), consistent with banks recording more charge-offs of homogeneous loans having more persistent future income before provisions for loan losses than other banks. In contrast, the coefficient on $X_t*NLCO(HET)_t$ is always negative, and significantly so at the three-year horizon.

The coefficient on X_t is significantly positive and generally in excess of 1, consistent with the sample banks' income before the provision for loan losses having high persistence during the 1990s boom. The coefficients on PLL_t and ΔNPA_t are either not significant or significantly negative. The coefficient on $HOM\%_t$ is not significant.

In summary, in this section we find that, during the 1990s boom, banks recording more charge-offs of homogeneous loans had higher and more persistent future income before provisions for loan losses over the next three years than other banks. These results are consistent with these banks smoothing income over a prolonged horizon.[17]

V. CONCLUSION

In this paper, we illustrate the contextual and asymmetric nature of income smoothing in the setting of banks' provisions for loan losses and loan charge-offs. Our evidence indicates that during the 1990s boom profitable banks smoothed their income downward by overstating provisions for loan losses for homogeneous loans, and they obscured this income smoothing by accelerating charge-offs of those loans and by recording more gross charge-offs when they inevitably recovered more charge-offs of these loans. Our results are in striking contrast with those of prior research that banks delayed provisions for loan losses primarily for heterogeneous loans during the pre-1990s bust period. Collectively, our results and those of prior research indicate that banks' exercise of discretion over provisions for loan losses and loan charge-offs is constrained by whether they hold homogeneous or heterogeneous loans, and that these constraints bind differently in the bust and boom phases of the business cycle.

While we examine a specific context, contextual and asymmetric income smoothing is likely to occur beyond the banking industry and as a result of firm-specific or industry-level rather than macroeconomic phenomena. We believe that research on income smoothing would benefit in terms of economic interest, faithfulness to the rules and practices of accounting, and statistical power by seeking out and documenting discretionary behavior that manifests itself in more contextual and less purely symmetrical ways than is captured in the usual research designs.

Our results also bear on the recent debate about rules- versus principles-based accounting standards. While rules-based standards are often thought to yield less discretionary accounting, our results suggest that an accounting rule that appears particularly nondiscretionary—the charging off of homogeneous loans more than a given number of days past due—leads to more discretion in certain contexts.

REFERENCES

Ahmed, A., C. Takeda, and S. Thomas. 1999. Bank loan loss provisions: A reexamination of capital management, earnings management and signaling effects. *Journal of Accounting and Economics* 28: 1–25.

---

[17] Consistent with our hypothesis and findings regarding income persistence, in prior versions of this paper we predict and find that $NLCO(HOM)_t$ is significantly positively associated with banks' contemporaneous April–March share returns during the 1990s boom, controlling for current income before the provision for loan losses and various indicators of loan default. We also predict and find that $NLCO(HOM)_t$ is significantly positively associated with banks' April–March share returns for the next year, controlling for current income before the provision for loan losses and various indicators of loan default, and that this association was fully explained by the association of $NLCO(HOM)_t$ with the next year's income before the provision for loan losses and recoveries. These results suggest that the market partially understands the discretionary behavior we document.

Bassett, W., and E. Zakrajsek. 2001. Profit and balance sheet developments at U.S. commercial banks in 2000. *Federal Reserve Bulletin* (June): 367–393.

Beatty, A., S. Chamberlain, and J. Magliolo. 1995. Managing financial reports of commercial banks: The influence of taxes, regulatory capital, and earnings. *Journal of Accounting Research* 33: 231–261.

———, B. Ke, and K. Petroni. 2002. Earnings management to avoid earnings declines across publicly and privately held banks. *The Accounting Review* 77: 547–570.

Beaver, W., C. Eger, S. Ryan, and M. Wolfson. 1989. financial reporting, supplemental disclosures, and bank share prices. *Journal of Accounting Research* 27: 157–178.

———, and E. Engel. 1996. Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics* 22: 177–206.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York, NY: John Wiley & Sons.

Berger, A., and G. Udell. 2002. Small business credit availability and relationship lending: The importance of bank organizational structure. *The Economic Journal* 112: F32–F53.

———, and ———. 2003. *The Institutional Memory Hypothesis and the Procyclicclity of Bank Lending Behavior.* Federal Reserve Board, Finance and Economics and Discussion Series 2003-4. Washington, DC: Federal Reserve Board.

Boehmer, E., and W. Megginson. 1990. Determinants of secondary market prices for developing country syndicated loans. *Journal of Finance* 45: 1517–1540.

Collins, J., D. Shackelford, and J. Wahlen. 1995. Bank differences in the coordination of regulatory capital, earnings, and taxes. *Journal of Accounting Research* 33: 263–291.

Dechow, P., and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14: 235–250.

Elliott, J., J. Hanna, and W. Shaw. 1991. The evaluation by the financial markets of changes in bank loan loss reserve levels. *The Accounting Review* 66: 847–861.

Federal Financial Institutions Examination Council. 1999. Uniform retail credit classification and account management policy. *Federal Register* 64: 6655–6659.

Federal Reserve Board. 2005. *Commercial Bank Examination Manual.* Washington, DC Federal Reserve Board.

Financial Accounting Standards Board (FASB). 1975. *Accounting for Contingencies.* Statement of Financial Accounting Standards No. 14. Norwalk, CT: FASB.

Galai, D., E. Sulganik, and Z. Wiener. 2003. Accounting values versus market values and earnings management in banks. Working paper, The Hebrew University of Jerusalem, Israel.

Grammatikos, T., and A Saunders. 1990. Additions to bank loan-loss reserves: Good news or bad news? *Journal of Monetary Economics* 25: 289–304.

Gray, R. 2004. The allowance for loan losses and earnings management. Working paper, University of Sydney.

Greenawalt, M., and J. Sinkey. 1988. Bank loan loss provisions and the income smoothing hypothesis: An empirical analysis, 1976–1984. *Journal of Financial Services Research* 1: 301–318.

Greenspan, A. 2001. *Condition of the U.S. Banking System.* Testimony of Federal Reserve Board Chairman before the U.S. Senate Committee on Banking, Housing, and Urban Affairs, June 20. Washington, D.C.: Government Printing Office.

Griffin, P., and S. Wallach. 1991. Latin American lending by major U.S. banks: The effects of disclosures about nonaccrual loans and loan loss provisions. *The Accounting Review* 66: 830–846.

Joyce, W. 1996. Management of loan loss reserves by commercial bankers—Part 1. *Journal of Bank Cost & Management Accounting* 9: 5–26.

Kanagaretnam, K., G. Lobo, and R. Mathieu. 2003. Managerial incentives for income smoothing through bank loan loss provisions. *Review of Quantitative Finance and Accounting* 20: 63–80.

Liu, C., and S. Ryan. 1995. The effect of bank loan portfolio composition on the market reaction to and anticipation of loan loss provisions. *Journal of Accounting Research* 33: 77–94.

———, ———, and J. Wahlen. 1997. Differential valuation implications of loan loss provisions across banks and fiscal quarters. *The Accounting Review* 72: 133–146.

Ma, C. 1988. Loan loss reserves and income smoothing: The experience in the U.S. banking industry. *Journal of Business, Finance, and Accounting* 15: 487–497.

McNichols, M., and G. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

Moyer, S. 1990. Capital adequacy regulations and accounting changes in commercial banks. *Journal of Accounting and Economics* 13: 123–154.

Musumeci, J., and J. Sinkey. 1990. The international debt crisis and bank loan-loss-reserve decisions: The signaling content of partially anticipated events. *Journal of Money, Credit, and Banking* 22: 370–387.

Office of the Comptroller of the Currency. 1993. *Interagency Policy Statement on the Allowance for Loan and Lease Losses*. Federal Deposit Insurance Corporation, Federal Reserve Board, and Office of Thrift Supervision. Washington, DC: Office of the Comptroller of the Currency.

———. 1998. *Allowance for Loan and Lease Losses: Comptroller's Handbook*. Washington, DC: Office of the Comptroller of the Currency.

Robb, S. 1998. The *Effect of Analysts' Forecasts on Earnings Management in Financial Institutions. Journal of Financial Research* 21: 315–331.

Ronn, E., and A. Verma. 1986. Pricing risk-adjusted deposit insurance: An options-based model. *Journal of Finance* 41: 871–895.

Ryan, S. 2002. *Financial Instruments and Institutions: Accounting and Disclosure Rules*. Hoboken, NJ: John Wiley & Sons.

Scheiner, J. 1981. Income smoothing: An analysis in the banking industry. *Journal of Bank Research* 12: 119–124.

Udell, G. 1989. Loan quality, commercial loan review and loan officer contracting. *Journal of Banking and Finance* 13: 367–382.

U.S. General Accounting Office (GAO). 1994. *Depository Institutions: Divergent Loan Loss Methods Undermine Usefulness of Financial Reports*. Washington, D.C.: Government Printing Office.

U.S. Securities and Exchange Commission (SEC). 1999a. *Sample Letter Sent to Some Bank Holding Companies*. Washington, D.C.: Government Printing Office.

———. 1999b. *Selected Loan Loss Allowance Methodology and Documentation Issues*. Staff Accounting Bulletin No. 102. Washington, D.C.: Government Printing Office.

Wahlen, J. 1994. The nature of the information in commercial bank loan loss disclosures. *The Accounting Review* 69: 455–478.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 40: 817–838.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 443–472

Classification and Market Pricing of the Cash Flows and Accruals on Trading Positions

Stephen G. Ryan
New York University

Jennifer W. Tucker
University of Florida

Paul A. Zarowin
New York University

ABSTRACT: Despite the classification of the cash flows on trading positions as operating under SFAS No. 102, trading is economically a hybrid operating/non-operating activity. Reflecting this hybrid nature, we hypothesize and find that the change in net trading assets has a less positive association with returns and future CFO than do the pure operating components of cash flows and accruals, and it has a more positive association with returns and future CFO than do the pure non-operating components of cash flows. Our paper is the first to propose and test hypotheses about the valuation implications of such hybrid cash flows and accruals.

Keywords: *trading; cash flows; accruals; classification; market pricing; banks.*

Data Availability: *All data are available from public sources.*

I. INTRODUCTION

In this paper we investigate whether the market prices the change in net trading assets (trading assets minus trading liabilities)[1] as an operating or non-operating activity or mixture of the two, and whether this market pricing is consistent with the association of the change in net trading assets with future cash flow from operations (CFO). Our investigation is motivated by the observation that—despite the classification of the cash

---

[1] Trading assets include trading securities, derivative assets not designated as accounting hedges, and other trading assets (e.g., commodity contracts). Trading liabilities include derivative liabilities not designated as accounting hedges and other trading liabilities. Throughout our sample period 1991–2003, banks were required to account for trading assets and liabilities at fair value under the AICPA Audit and Accounting Guides for Banks and Savings Institutions and for Brokers' and Dealers in Securities or other accounting standards (e.g., SFAS No. 115, FASB 1993).

The authors thank Sang-Hyun Suh for able research assistance and two anonymous referees, John Dickhaut, Frank Gigler, and seminar participants at the University of Minnesota for useful comments.

Editor's note: This paper was accepted by Patricia Dechow, Editor.

Submitted May 2004
Accepted October 2005

flows on trading positions as operating under SFAS No. 102—trading is economically a hybrid operating/non-operating activity. This is because the change in net trading assets equals the net unrealized gain, a primarily operating item, plus the net principal cash outflow, a primarily non-operating item, on trading positions during the period. Our hypotheses and tests are based on the maintained assumption, supported by theory and empirical research, that operating activities have stronger implications for returns and future CFO than do non-operating activities.

The operating aspects of trading pertain to the generation of trading revenue, which includes trading gains and losses as well as fees (e.g., commissions or spreads) for trading-related activities such as dealing. Accounting standards focus on these aspects of trading. For example, SFAS No. 115 (FASB 1993) defines trading as "active and frequent buying and selling ... with the objective of generating profits [i.e., gains] on short-term differences in price." SFAS No. 119 (FASB 1994) includes dealing under its definition of trading. The non-operating aspects of trading pertain to the principal cash flows at the inception of investments in trading assets or fundraising through trading liabilities and to the subsequent return or repayment of that principal. In practice, however, the operating and non-operating aspects of trading are inseparable, because the generation of trading revenue often requires traders to assume trading positions with initial principal (i.e., trading assets and liabilities with non-zero principal amounts at inception).

Based on the accounting literature, financial economic theories, and empirical findings summarized in Sections II and III, we expect that share returns and future CFO are positively associated with current CFO and accruals, because the operating activities that give rise to CFO and accruals tend to be positive present value and persistent. In contrast, we expect that returns and future CFO have a near zero association with cash flows from investing (CFI) and financing (CFF), because the non-operating activities that give rise to CFI and CFF tend to be close to zero net present value. Reflecting the hybrid nature of trading, we hypothesize and find that the change in net trading assets has a less positive association with returns and future CFO than do the pure operating components of CFO and accruals and a more positive association with returns and CFO than do CFI and CFF. Our paper is the first to propose and test hypotheses about the valuation implications of such hybrid cash flows and accruals.

While our research question applies to all firms engaging in trading, we examine a sample of 37 U.S. banks that hold appreciable amounts of trading assets. We chose this sample because it is the largest available set of fairly homogeneous, publicly traded firms with appreciable trading positions.

We conduct three additional empirical analyses to investigate further the valuation implications of hybrid operating/non-operating flows. First, as a benchmark for our results regarding trading positions, we examine the associations between both returns and future CFO and the change in held-for-sale (HFS) loans. HFS loans constitute the closest and economically most important analogue to trading positions for our banking sample. Like trading, originating HFS loans is a hybrid operating and non-operating activity. We expect the non-operating aspects of HFS loans to be relatively more important than for trading, however, because most HFS loans are residential mortgages or consumer loans that are originated and sold in highly competitive markets, rendering this activity closer to zero present value. We hypothesize and find that returns and future CFO are less positively associated with the change in HFS loans than with the change in net trading assets.

Second, in our primary empirical analysis described above, we examine the valuation implications of the total change in net trading assets, not its primarily operating (net unrealized gain) and non-operating (principal cash flow) components. We do this in part because we cannot separately observe unrealized gains and the principal cash flows on trading positions and in part because of the inseparability of the operating and non-operating aspects of trading. We are able to observe trading revenue for a subset of observations, however, and for these observations we use this variable as a proxy for net unrealized gains on net trading assets during the period. We hypothesize and find that returns and future CFO are positively associated with trading revenue controlling for the change in net trading assets.

Third, to address potential heterogeneity in our sample and to examine whether the market differentiates banks based on the relative operating nature of their activities, we decompose the sample into six large derivatives dealers with both trading assets and liabilities and 31 non-dealer banks with only trading assets. Due to the dealers' greater focus on fee generation, we expect their activities to be more operating than those of non-dealer banks. We hypothesize and find that returns and future CFO are more positively associated with the pure operating components of CFO and accruals and with the change in HFS loans for dealers than for non-dealer banks. Unexpectedly, we find that the change in net trading assets exhibits no differential market pricing and only a marginally significantly different association with future CFO for dealers and non-dealer banks.

Our results generally support the conclusions that the market appreciates the hybrid nature of the cash flows and accruals associated with trading positions and HFS loans and that it differentiates banks based on the relative operating character of their activities. This likely reflects the fact that these items are disclosed separately from other cash flows and accruals in banks' financial reports, and so investors can adjust their valuation analyses for these items. Our results imply that it is important that the cash flows and accruals for hybrid activities be reported in a disaggregated fashion on the statement of cash flows (SCF) and elsewhere in financial reports, regardless of how these items are classified.

Our results bear on two related, important, and timely financial reporting issues. First, the proper classification of cash flows on trading positions is of current interest due to recent highly publicized cases in which Enron and other trading firms issued liabilities that provided them with cash at inception—prepaid derivatives or commodity contracts (prepays)—that they classified as trading and for which they classified the initial principal cash flow in CFO. In a prepay, one firm receives cash at the inception of the contract in exchange for the promise of future delivery of cash, commodities, or other assets to its counterparty. As such, prepays involve financing.

Addressing the cash flow classification issues raised by prepays, SFAS No. 149 (FASB 2003) requires that if a derivative contains "an other-than-insignificant financing element ... at inception, then the borrower shall report all cash inflows and outflows associated with that derivative" as financing. It seems likely that SFAS No. 149's classification requirement will be applied by analogy to all trading liabilities, prepays in particular, regardless of whether they are deemed derivatives for accounting purposes. This classification is in contrast to the operating classification of the principal cash flows on trading liabilities during our sample period. Our analysis and results imply that the classification of the hybrid cash flows on trading positions as either operating or non-operating is limited. We discuss possible remedies for this problem in the conclusion.

Second, the distinctions between operating and non-operating activities and thus between operating and non-operating cash flows made in SFAS No. 95 (FASB 1987) are inherently blurry for the large and increasingly economically important set of trading firms and financial institutions whose operations primarily involve investing and financing.[2] For such firms, operating and non-operating activities are better viewed as ends of a continuum, with different types of activities falling at different places along the continuum. In this study, we focus on the market's perception of where trading, which SFAS No. 102 views as an operating activity, falls along this continuum. We adopt this focus because in our view trading constitutes the most salient example of the phenomena we wish to describe and explain.[3]

The remainder of the paper is organized as follows. Section II summarizes the related research literatures and our contributions to them. Section III discusses GAAP and accounting practices regarding SCF classification for trading positions and HFS loans during our sample period, and it motivates and states our hypotheses. Section IV describes our sample, defines our cash flow and accrual components, and discusses the descriptive statistics. Section V develops our returns and future CFO regression equations, restates our hypotheses as coefficient restrictions in those equations, and reports the estimation results. Section VI reports the results of various specification analyses. Section VII concludes.

II. PRIOR RESEARCH

Our paper contributes to four distinct literatures in accounting research: (1) on the market pricing of (the components of) cash flows and either earnings or accruals for nonfinancial firms, (2) on the market pricing of the components of earnings for financial institutions, (3) on the distinct nature of trading firms and the usefulness of the financial information they provide, and (4) on the importance of the financial statement classification of hybrid items. We discuss each of these literatures and our contributions to them in turn.

The first literature—examining the market pricing of (the components of) cash flow and either earnings or accruals for nonfinancial firms—is large and well established.[4] Of this literature, Livnat and Zarowin (1990) is the paper most closely related to ours, being the only one to examine the valuation implications of the components of cash flow. For a sample of nonfinancial firms, Livnat and Zarowin (1990) predict based on financial economic theories and find that the market prices CFO, CFI, CFF, and accruals differently. Specifically, they find that returns are significantly positively associated with both CFO and accruals, less strongly and significantly negatively associated with CFI, and insignificantly positively associated with CFF. Consistent with Livnat and Zarowin's (1990) finding that returns are positively associated with CFO and accruals, for samples of nonfinancial firms, Dechow et al. (1998) and Barth et al. (2001) find that CFO and the components of accruals are significantly positively associated with future CFO. We contribute to this literature by developing and testing hypotheses about the implications for returns and future CFO of the hybrid trading and HFS loan components of cash flows and accruals. We expect these components to fall at different places along the continuum between the primarily operating and non-operating items examined by the prior literature.

The second literature—examining the valuation implications of components of earnings for financial institutions—is also large and well established. For example, for samples of property-casualty insurers, Foster (1977) examines the market pricing of the underwriting, investment, and capital gains components of earnings, and Beaver and McNichols (2001) examine the market pricing of the CFO, accrual, and loss reserve development components of earnings. For samples of banks, Barth et al. (1990) examine the market pricing of earnings before security gains and losses and realized security gains and losses, Barth (1994) examines the market pricing of these variables and also unrealized security gains and losses, and Wahlen (1994) examines the market pricing and future CFO implications of earnings before the provision for loan losses, the provision for loan losses, and various other loan default variables. While some of these prior studies examine gains and losses and find them to be value-relevant, they do not focus on trading gains and losses, as we do. We contribute to this literature by being the first study to examine the market pricing of components of cash flows for any type of financial institution. The distinctions between operating and non-operating activities and thus between operating and non-operating cash flows are inherently blurry for financial institutions, because their operations involve investing and financing.

---

[2] See Ryan (2002) for elaboration of this point in the contexts of various types of financial institutions.

[3] Our study may also have implications for the analogous SCF classification issues that arise for nonfinancial firms as well. For example, many academics and practitioners (e.g., Nurnberg 1993) have criticized nonfinancial firms' classification of the following cash flows as operating under SFAS No. 95 and SFAS No. 102 (FASB 1987, 1989): interest receipts and expenditures, dividend receipts, expenditures for inventory to stock newly opened stores, and installment sales receipts. Of these cases, installment sales, which involve both operations and financing, exhibit a nature closest to the hybrid nature of trading.

[4] See, for example, Wilson (1986, 1987), Bowen et al. (1987), Rayburn (1986), Bernard and Stober (1989), Ali (1994), and Dechow (1994).

The third literature—examining the distinct nature of trading firms and the usefulness of the financial information they provide—is both small and recent. Jorion (2002) and Liu et al. (2004) examine the risk-relevance of banks' Value-at-Risk disclosures for their trading portfolios. These papers relate to our study insofar as they are affected by some of the same economic and financial reporting phenomena that motivate our analysis, for example, banks' matching of trading assets and liabilities. We contribute to this literature by providing new insights into the distinct nature of trading firms, in particular, how their trading positions blur the distinctions between operating and non-operating activities.

The fourth literature—examining the importance of financial statement classification of hybrid items—is also fairly small and recent. In an experimental study, Hopkins (1996) shows that buy-side financial analysts' assessment of a specific financing instrument with a hybrid nature—mandatorily redeemable preferred stock—is affected by how that instrument is classified on the balance sheet. Hopkins' (1996) study is motivated by psychology/judgment and decision-making research regarding how knowledgeable individuals organize and interpret information.[5] Cheng et al. (2000) and Linsmeier et al. (2000) empirically examine the association between market value and various hybrid financing instruments that may be classified as debt, equity, or in the mezzanine on the balance sheet. Our paper contributes to this empirical literature by being the first market-based study of the relation between the market pricing and SCF classification of hybrid items.

III. RELEVANT GAAP AND HYPOTHESIS DEVELOPMENT

Relevant GAAP

GAAP governing the classification of cash flows acknowledge that certain activities have both operating and non-operating aspects. SFAS No. 95 provides intuitive and well-known characterizations of investing and financing activities that we do not repeat here. Rather than defining operating activities directly, SFAS No. 95 (paragraph 21) defines operating activities as a catchall category that includes "all transactions and other events that are not defined as investing or financing activities." This suggests that hybrid transactions that do not cleanly fall into the investing or financing categories should be classified as operating. However, in SFAS No. 95 (paragraph 24), FASB states that "[c]ertain cash receipts and payments may have aspects of more than one class of cash flows ... [i]f so, the appropriate classification shall depend on the activity that is to be the predominant source of cash flows for the item."[6]

---

[5] John Dickhaut points out that classification can affect market prices even under models with fully rational investors depending on the specific process by which prices form, and so psychology theories are not the only possible reason why classification might matter. We believe his point is worthy of future research by accounting theorists, who up to this point have not focused on the effects of financial statement classification.

In SFAS No. 102, FASB (1989) decided that the predominant characteristics of trading positions and HFS loans are operating, requiring in SFAS No. 102 (paragraphs 8 and 9) that the cash flows on assets "acquired specifically for resale and ... carried at market value in a trading account" and loans "originated or purchased specifically for resale and ... held for short periods of time" be classified as operating. In SFAS No. 102 (paragraph 26), FASB justified this requirement stating that these assets are "similar to inventory in other businesses." No GAAP rule developed prior to the issuance of SFAS No. 149 in April 2003 specifically considers the classification of the cash flows from trading liabilities with initial principal or any other type of liability that has both operating and financing aspects. Prior to SFAS No. 149, the practice for our sample of banks (and, in our understanding, for trading firms generally) was to classify the cash flows on trading liabilities as operating, consistent with the classification of the cash flows on trading assets. As discussed in the introduction, this practice likely will change due to the issuance of SFAS No. 149.

Hypothesis Development

We expect the associations of returns and future CFO with the cash flows and accruals resulting from an activity to rise with the extent to which that activity is positive present value (or is associated with other positive present value activities), and, if the activity is positive present value, with its persistence. We expect operating activities to be more positive present value and possibly more persistent than are non-operating activities for the following reasons. First, capital markets are relatively efficient compared to operating markets and non-market settings. Second, numerous financial economic theories surveyed by Livnat and Zarowin (1990) imply that managerial inside information and other market imperfections can cause investing and financing activities to be either positive or negative present value, with these effects possibly offsetting. Third, an August 1999 G4+1 discussion paper on reporting financial performance concludes that operating activities tend to be "value adding" and "recurring." Fourth, Livnat and Zarowin (1990) find that returns are significantly positively associated with CFO and accruals, are less strongly and significantly negatively associated with CFI, and are insignificantly positively associated with CFF. Finally, Dechow et al. (1998) and Barth et al. (2001) find that future CFO is significantly positively associated with CFO and the components of accruals.

For banks, operating activities often generate related ongoing streams of fee income, may involve skill or other comparative advantage that yields positive net present value, and tend to be highly repetitive. Operating activities are especially likely to be positive present value for large banks that engage in investment banking and other high-margin services businesses. In contrast, banks' non-operating activities are less likely to be associated with ongoing streams of fee income and are more likely to be close to zero net present value, although they often are highly persistent. In particular, we expect the value implications of investing and financing activities predicted by the financial economic theories mentioned above to be relatively small (though not necessarily zero) for banks, due to the highly competitive nature and frequency of those activities. For these reasons, we expect banks' returns and future CFO to be more positively associated with their CFO and accruals than with their non-operating cash flows.

---

[6] In paragraphs 93–95 of SFAS No. 95, FASB provides the example of installment sales—which have both operating and investing characteristics for the seller—to illustrate how the predominant characteristics should be determined. The FASB concludes that the predominant characteristic of installment sales is operating, stating "cumulative cash flow from operating activities over the life of an enterprise that finances most of its sales under installment plans might be negative." In reaching this conclusion, FASB appears to be primarily concerned with the usefulness of CFO as a performance measure.

Trading has characteristics of an operating activity insofar as it requires managerial effort and skill to identify mispriced assets and liabilities and to manage complex and interrelated positions and this effort and skill generates positive net present value through trading gains or fee income for services such as dealing, financial advice, and securities custody and processing. Though typically held for a shorter time than non-trading assets and liabilities, trading assets (liabilities) constitute investments (provide financing) for as long as they are held. Thus, trading has both operating and non-operating aspects. These aspects are inseparable, however, because the generation of trading revenue often requires traders to take positions with initial principal. For example, a derivatives dealer that writes a call option for a client receives a cash premium at inception; the portion of that premium that exceeds the fair value of the option is trading revenue and the remainder of the premium is the initial principal amount of a trading liability.[7]

Like trading positions, HFS loans have characteristics of an operating activity insofar as originating these loans requires managerial effort and skill to solicit and screen borrowers and this effort and skill generate positive net present value or are associated with services that yield fee income (e.g., loan origination and servicing). Though typically held for a shorter time than loans held in portfolio, HFS loans constitute investments for as long as they are held.

Trading positions have relatively more of an operating nature than HFS loans, however, for two reasons. First, the management of trading positions both individually and at the portfolio level can be complex, and these positions are more likely to have positive net present value and to be associated with high-margin services. In contrast, most HFS loans are residential mortgages or consumer loans, commodity financial products that are originated, serviced, and sold in highly competitive markets. Second, many types of trading firms (e.g., dealers and arbitrageurs) manage their trading positions at the portfolio level, assuming largely matched positions in trading assets and liabilities, where their trading liabilities both finance and hedge their trading assets. Well-matched trading portfolios require little net investment or provide little net financing, thereby accentuating the operating aspects of those portfolios.

In summary, trading positions and HFS loans are mixtures of operating and non-operating activities, with the mix being tilted relatively more toward operating for trading positions than for HFS loans. Under the maintained assumption that investors have some (though not necessarily a perfect) understanding of these classification issues, we hypothesize (in alternative form) that the associations of both returns and future CFO with changes in net trading assets and HFS loans reflect these mixtures as follows:

H: Returns and future CFO have:

- relatively high positive associations with the portion of operating cash flows and accruals arising from activities other than trading and HFS loans,

---

[7] While GAAP have changed somewhat over time, dealers generally are allowed to recognize trading revenue (dealer profit) at the inception of trading positions if the market provides sufficient evidence of fair value (see AICPA Audit and Accounting Guide for Brokers and Dealers in Securities, EITF 00-17, and EITF 02-3).

- a lower positive association with the change in net trading assets,
- an even lower positive association with the change in HFS loans, and
- relatively small positive or negative associations with CFI and CFF.

This compound hypothesis contains four separate hypotheses about the levels of the associations of returns and future cash flows with specific components of cash flows and accruals. In Sections V and VI we restate Hypothesis H as restrictions on differences between coefficients in our regression equations.

IV. SAMPLE, DEFINITIONS, AND DESCRIPTIVE STATISTICS

Sample

To be included in our sample, a firm must be a U.S. commercial bank or thrift (hereafter bank) appearing on the 2002 or 2003 Annual Bank Compustat file that has at least $500 million in total assets and trading assets equal to at least 1 percent of total assets in at least one year in our sample period. Since banks' regulatory capital leverage ratios[8] average 8 to 9 percent during our sample period, the sample banks' trading assets typically exceed that percentage of their owners' equity. We imposed the conditions on total and trading assets to ensure that the banks had appreciable trading operations. We did not require these conditions to be met in each year for a given bank in order to preserve time-series data and thus statistical power. We did not include Citigroup after the 1998 acquisition of Citicorp by Travelers in the sample due to the concern that banks and insurance companies' cash flows and accruals would have different natures and thus different valuation implications.

To the extent possible, we obtained financial report data from Bank Compustat. Since Bank Compustat does not include SCF data, we manually collected the following variables from the SCFs for each bank for every year the variables were available online on the SEC's EDGAR database: *CFO*, *CFI*, *CFF*, the change in net trading assets, the change in *HFS* loans, and changes in all other accrual accounts during the year. Because the EDGAR database begins with 1993 filings by a subset of firms and SCFs include three years of data, our SCF data begins in 1991 with the seven banks that filed on EDGAR in 1993.[9] For 27 (8 percent) of our observations, banks did not provide separate line items for the change in net trading assets on their SCF, and in these cases we constructed these line items using the changes in these balance sheet amounts if the banks provided this information in two consecutive years. None of our results change substantively if we drop these observations.

We obtained annual returns from May of the fiscal year to April following the fiscal year, adjusted for stock splits and dividends, denoted R, from the 2003 CRSP Monthly file. Combining the data from Bank Compustat, EDGAR, and CRSP yields 330 observations from 1991–2003 with the complete data needed for the main returns regressions analyses. Table 1, Panel A reports the effect of each of the above conditions on the sample size.

The banks in our sample are individually large, and they hold a high and increasing percentage of the assets of the U.S. banking system. For example, these banks held 31

---

[8] Banks' leverage ratio for regulatory capital purposes is Tier 1 capital divided by regulatory assets, which roughly corresponds to owners' equity divided by assets.

[9] For certain observations, the data taken from the 1993 filings on EDGAR for the 1991 and 1992 fiscal years may be restated for mergers accounted for as poolings of interest. In Section VI, we discuss the results of a specification test in which observations that appear to have had significant mergers or acquisitions or significantly restated earnings in a year are eliminated. This specification test yields results substantively similar to those reported in this paper.

TABLE 1
Sample

Panel A: Sample Construction

U.S. banks on 2002 or 2003 Bank Compustat	734
− U.S. banks with less than $500 million assets	644
= U.S. banks that with more than $500 million assets	90
− U.S. banks with trading assets less than 1 percent of total assets	(53)
= Sample banks	37

Panel B: Sample Banks[a, b]

	Total Assets	Trading%	*CFO*	*ACC*	*ΔTA*	*ΔHFS*	*CFI*	*CFF*
Derivatives dealers[c]								
Bank of America	392.6	.068	.157	−.013	.073	.022	−.302	.156
Bank of New York	62.4	.055	.102	−.010	.008	.096	.156	−.219
Citicorp	261.3	.105	.097	.022	.247	.011	−.563	.453
Fleetboston Financial	136.0	.017	.192	−.089	.015	.028	.244	−.475
JPMorgan Chase & Co	408.9	.186	.142	.020	.178	.000	−.710	.551
Wachovia Corp	191.7	.043	.186	−.063	.038	.032	.108	−.223
Non-dealer banks								
Bank One Corp	210.2	.030	.204	−.112	−.037	.000	−.274	.048
Bankatlantic Bancorp[d]	4.4	.012	.549	−.458	.068	−.281	−1.691	1.425
Beverly Hills Bancorp[d]	1.4	.012	−1.814	1.832	−.006	1.920	−.531	1.277
Chester Vy Bancorp	0.6	.000	.007	.010	.005	.000	.122	.017
City National Corp	6.6	.009	.147	−.100	−.009	.000	−.104	−.175
Commerce Bancorp	8.2	.011	.140	−.051	.022	−.025	−1.594	1.536
Community West Bancshares	0.3	.020	−.129	.168	−.051	.443	.149	.353
First Defiance Financial[d]	1.1	.010	−1.027	.859	−.046	.529	−.072	1.179
First Horizon National Corp	8.8	.017	−.056	.206	.052	.270	−.749	.717
First Tennessee National	16.7	.019	−.005	.090	.017	.073	−.268	.294
Irwin Financial Corp	2.6	.027	−.386	.454	.040	.485	−.856	1.274
M & T Bank Corp	25.0	.006	.115	−.027	.002	.029	−.365	.280
Massbank Corp	0.9	.047	.058	.029	.019	.005	.116	.014
Mellon Financial Corp	40.7	.006	.148	−.044	.000	.015	.013	−.130
National Commerce Financial	11.0	.008	.073	−.009	.008	.001	−.538	.462
National Penn Bancshares	2.0	.006	.072	−.009	.008	−.002	−.622	.557
Ocwen Financial Corp[d]	2.3	.056	.095	−.131	−.188	.006	.378	−.466
Oriental Financial Group	1.9	.017	.018	.093	.001	.077	−1.078	1.051
Popular Inc	20.9	.009	.202	−.085	.012	.030	−.896	.701
R & G Financial Corp	4.2	.077	−.363	.607	.289	.625	−3.782	4.072
Regions Financial Corp	27.6	.005	.114	−.013	.002	.018	−.437	.353
Republic Bancshares Inc	2.0	.010	.077	−.054	.074	.000	−1.272	1.209
State Street Corp	45.2	.012	.079	−.011	.010	.000	−1.022	.953
Sterling Bancshares	2.9	.034	.035	.034	.085	−.032	−.299	.266
Suntrust Banks Inc	58.2	.003	.101	−.017	−.004	.015	−.406	.305
Union Planters Corp	26.8	.010	.219	−.083	.014	.000	−.117	−.032
Unionbancal Corp	33.2	.009	.162	−.037	.014	−.001	−.658	.539
Webster Financial Corp[d]	8.1	.007	.201	−.041	.029	.043	−.963	.369
Wells Fargo & Co	167.2	.010	.080	.016	.009	.097	−.328	.201
Woronoco Bancorp[d]	0.8	.012	.218	−.141	−.121	.007	−1.025	.801
Zions Bancorporation	13.6	.016	.081	.015	.014	.025	−.407	.355

(*continued on next page*)

TABLE 1 (Continued)

[a] *Trading%* denotes the ratio of trading assets to total assets; *CFO* denotes cash flow from operations; *ACC* denotes total accruals, i.e., the difference between net income and *CFO*; ΔTA denotes the change in net trading assets; ΔHFS denotes the change in held-for-sale loans; *CFI* denotes cash flow from investing; *CFF* denotes cash flow from financing. *Total Assets* are in billions. All variables other than *Total Assets* and *Trading%* are deflated by beginning-of-year market value of equity.

[b] ΔHFS is assumed to be zero for bank-years it is not reported on the Statement of Cash Flows.

[c] Indicates derivatives dealers that report changes in both trading assets and trading liabilities in their Statement of Cash Flows.

[d] Indicates thrifts.

percent (25 percent) of the assets of U.S. commercial banks (commercial banks and thrifts) in 1995 and 45 percent (38 percent) of the assets of U.S. commercial banks (commercial banks and thrifts) in 2003.[10] Mergers and acquisitions are the primary cause of the rise in these percentages over time. Ignoring our omission of Citigroup after its acquisition by Travelers in 1998, our sample banks held almost all of the trading assets and trading liabilities of the U.S. banking system throughout our sample period.

Our sample contains two distinct subsamples. Six of the largest commercial banks—Bank of America, Bank of New York, Citicorp, FleetBoston, JPMorgan Chase, and Wachovia—collectively dominate the global over-the-counter derivatives-dealing markets. Derivatives dealers closely match the value and risk characteristics of their derivatives assets and liabilities. As a consequence of this matching, the six dealers have substantial trading liabilities that average about half the amount of their trading assets on average, with the excess of their trading assets over their trading liabilities reflecting their classification of a significant portion of their investment securities as trading. The remaining 31 non-dealer banks have minimal or no trading liabilities, and their trading assets are primarily securities. To determine the nature and extent of the heterogeneity across the two subsamples, in addition to testing our hypotheses on the full sample, we test for differences across the two subsamples.

Table 1, Panel B lists the 37 banks (31 commercial banks and six thrifts) in the sample, distinguishing the dealers from the non-dealer banks. To provide a sense for the heterogeneity of the sample banks, the panel reports the mean values of total assets, trading assets as a percentage of total assets, and certain key regression variables defined below for each bank across the years it appears in the sample.

Definitions of the Cash Flow and Accrual Components

The banks in our sample all report CFO using the indirect method. This method begins with net income, denoted *NI*, and then subtracts certain items that for simplicity (but not entirely accurately, as discussed below) we refer to as total accruals, denoted *ACC*, in order to arrive at *CFO*. We decompose *ACC* into the change in net trading assets (trading assets minus trading liabilities), denoted ΔTA, the change in HFS loans, denoted ΔHFS, and all other accruals, which we refer to as ordinary accruals, denoted *OACC*. This breakdown reflects the following representation of the operating section of the SCF:

---

[10] Assets for the commercial banking industry are obtained from Carlson and Perli (2004) and assets for the thrift industry are obtained from Office of Thrift Supervision (2004).

$$NI - ACC = NI - OACC - \Delta TA - \Delta HFS = CFO. \tag{1}$$

Of the three components of *ACC*, only *OACC* reflects pure accruals, i.e., timing differences between *NI* and *CFO*. While ΔTA and ΔHFS include unrealized gains and losses, which are timing differences, they also include principal cash flows that never affect *NI* but do affect *CFO*. For example, ΔTA equals a timing difference—the unrealized gain on trading positions during the period—plus a net principal cash outflow—the purchases of net trading assets during the period minus the principal payments on net trading assets during the period minus the cost of net trading assets sold during the period. If we ignore the possibility of default by the firm or its counterparty, then the principal cash flows on ΔTA and ΔHFS sum to zero over the lives of those items.

Reworking Equation (1) slightly, we define ordinary cash flow from operations, *OCFO*, as *CFO* plus ΔTA and ΔHFS, which equals *NI* minus *OACC*:

$$OCFO = CFO + \Delta TA + \Delta HFS = NI - OACC. \tag{2}$$

Just as *OACC* reflects pure accruals, *OCFO* reflects pure operating cash flows.

The definitions in Equations (1) and (2) correspond to the following analytical reworking of the operating section of the SCF:

$$\begin{array}{l} NI \\ \underline{-\ OACC} \\ =OCFO \\ -\ \Delta TA \\ \underline{-\ \Delta HFS} \\ =CFO \end{array}$$

In the rest of this paper we refer to *OCFO* and *OACC* as the pure operating components of cash flow and accruals, respectively. We refer to ΔTA and ΔHFS as hybrid operating/non-operating components of cash flow and accruals, where as discussed above ΔTA is relatively more operating than ΔHFS. We refer to *CFI* and *CFF* as the pure non-operating components of cash flow.

To illustrate these definitions, the Appendix reports M&T Bank Corporation's 2000 SCF with the operating section of that statement reworked in the fashion described above. For this observation, both ΔTA (\$6.9 million) and ΔHFS (\$81.5 million) are positive, with ΔHFS being quite large, so that *OCFO* (\$435.1 million) is 26 percent more positive than reported *CFO* (\$346.6 million), and *OACC* (−\$148.9 million) is 146 percent more negative than *ACC* (−\$60.5 million). As indicated by the descriptive statistics for these variables discussed below, large differences between the pure operating and reported measures of *CFO* and accruals are common in our sample. The investing and financing sections of this SCF provide representative examples of non-operating cash flows for the banks in the sample that indicate the arbitrary nature of SCF classifications for banks and other financial institutions. For example, M&T Bank's cash flows from investing activities include the principal cash flows on (available-for-sale and held-to-maturity) investment securities that would have been classified in cash flows from operating activities had M&T Bank classified those securities as trading.

As discussed above, ΔTA equals unrealized gains (more operating) plus principal cash outflows for trading positions (less operating) during the period. Neither of these components of ΔTA is directly observable. In specification analysis reported in Section VI, we use the best available proxy for one of the components of ΔTA to distinguish the components and thereby analyze their distinct valuation consequences. Specifically, for the observations for which trading revenue, denoted *TR*, is reported, we use *TR* as a proxy for the unrealized gain component, so that $\Delta TA - TR$ proxies for the principal cash outflow component. *TR* includes unrealized and realized gains and may include some fee revenue from trading. As a result, *TR* overstates (understates) unrealized gains by the amount of realized gains (losses), and it overstates unrealized gains by the amount of fee revenue it includes. Since *TR* measures unrealized gains with error, $\Delta TA - TR$ measures the principal cash outflows with an error of the same magnitude but opposite sign. The amount of this error cannot be determined.

Descriptive Statistics

Table 2 reports descriptive statistics for the regression variables, total assets, and the ratio of trading assets to total assets. Panel A reports these statistics for the full sample, while Panels B and C report them for the subsamples of six dealers and 31 non-dealer banks, respectively. All of the regression variables other than *R* (which is naturally deflated) are deflated by the market value of equity at the beginning of the fiscal year. To mitigate the effect of outliers in the regression analyses, all of the regression variables are winsorized at three standard deviations from their means and descriptive statistics are reported on the winsorized variables.

Panel A of Table 2 shows that, like the rest of the banking industry, the sample banks generally performed well from 1991–2003, as evidenced by positive mean *R* (.248), *CFO* (.080), and *ACC* (.016). Moreover, the mean of *OCFO* (.177) is more than twice as large as the mean of *CFO* (.080), because the sample banks grew their net trading assets and HFS loans over the sample period (mean ΔTA = .028 and mean ΔHFS = .069), depressing CFO substantially. The standard deviations of ΔTA and ΔHFS, .220 and .339, are each substantial relative to the standard deviations of *CFO* and *ACC*, .438 and .413, suggesting that *CFO* and *ACC* are heterogeneously noisy measures of the pure operating components *OCFO* and *OACC*, respectively. Consistent with their good performance, the sample banks raised substantial financing (mean *CFF* = .443) and made significant investments (mean *CFI* = –.517).

Comparison of Panels B and C of Table 2 shows that, as expected, the six dealers have much higher mean total assets ($264 billion versus $32 billion), mean ratio of trading assets to total assets (.090 versus .017), and mean *TR* (.082 versus .023) than the 31 non-dealer banks. The dealers have much higher mean *CFO* (.149 versus .060) and *OCFO* (.265 versus .152) than the non-dealer banks, suggesting better performance, although mean *R* is similar for the two subsamples. The dealers grew their trading assets faster (mean ΔTA = .083 versus .013) but their HFS loans slower (mean ΔHFS = .033 versus .079) than the non-dealer banks.

Table 3 reports both Pearson and Spearman correlations for the variables in our returns regressions, many of which are significantly correlated. Most of the Pearson and Spearman correlations are similar, although we discuss one notable difference below. We summarize only the more interesting of these correlations and the underlying phenomena they suggest. *R* is more highly positively correlated with *OCFO* (Pearson = .176; Spearman = .262) than with *CFO* (Pearson = .059; Spearman = .070) and with *ACC* (Pearson = .008;

TABLE 2
Descriptive Statistics[a, b]

Variable	Mean	Std. Dev.	Minimum	Median	Maximum
Panel A: Full Sample (330 bank-years)					
R	0.248	0.404	−0.899	0.210	1.531
CFO	0.080	0.438	−1.816	0.105	1.929
$FCFO3^{d}$	0.609	0.955	−2.911	0.447	4.257
ACC	0.016	0.413	−1.929	−0.015	2.659
ΔTA	0.028	0.220	−1.025	0.003	1.108
ΔHFS^{c}	0.069	0.339	−1.646	0.000	1.925
OCFO	0.177	0.246	−1.068	0.122	1.387
OACC	−0.081	0.189	−0.793	−0.039	0.628
CFI	−0.517	1.082	−4.288	−0.378	3.232
CFF	0.443	1.137	−3.427	0.299	4.368
TR^{e}	0.038	0.070	−0.069	0.012	0.375
Trading%	0.031	0.052	0.000	0.013	0.328
Total assets	76,724	143,142	252.03	19,518	770,912
Panel B: Dealer Banks (73 bank-years)					
R	0.267	0.363	−0.450	0.238	1.463
CFO	0.149	0.349	−0.777	0.139	1.929
$FCFO3^{d}$	1.026	1.186	−1.058	0.649	4.257
ACC	−0.025	0.339	−1.819	−0.018	0.680
ΔTA	0.083	0.314	−1.025	0.018	1.108
ΔHFS^{c}	0.033	0.149	−0.607	0.000	0.774
OCFO	0.265	0.262	−0.155	0.191	1.387
OACC	−0.141	0.227	−0.793	−0.089	0.184
CFI	−0.151	0.960	−2.172	−0.288	3.232
CFF	0.012	0.968	−3.427	0.106	2.018
TR^{e}	0.082	0.101	0.008	0.039	0.375
Trading%	0.090	0.077	0.002	0.077	0.328
Total assets	263,564	216,569	39,426	200,500	770,912
Panel C: Non-Dealer Banks (257 bank-years)					
R	0.242	0.415	−0.899	0.205	1.531
CFO	0.060	0.459	−1.816	0.094	1.929
$FCFO3^{d}$	0.476	0.829	−2.911	0.419	2.969
ACC	0.028	0.431	−1.929	−0.012	2.659
ΔTA	0.013	0.182	−1.025	0.002	1.108
ΔHFS^{c}	0.079	0.375	−1.646	0.000	1.925
OCFO	0.152	0.236	−1.068	0.117	1.382
OACC	−0.064	0.173	−0.793	−0.031	0.628
CFI	−0.620	1.094	−4.288	−0.419	3.232
CFF	0.565	1.153	−3.427	0.377	4.368
TR^{e}	0.023	0.048	−0.069	0.007	0.262
Trading%	0.017	0.030	0.000	0.010	0.265
Total assets	31,703	61,706	252	11,857	387,798

(*continued on next page*)

TABLE 2 (Continued)

[a] R denotes share return for the 12 months ending on April 30 following the fiscal year-end; CFO denotes cash flow from operations; $FCFO3$ denotes the sum of the next three years' CFO; ACC denotes total accruals, i.e., the difference between net income and CFO; ΔTA denotes the change in net trading assets; ΔHFS denotes the change in held-for-sale loans; $OCFO$ denotes $CFO + \Delta TA + \Delta HFS$; $OACC$ denotes $ACC - \Delta TA - \Delta HFS$; CFI denotes cash flow from investing; CFF denotes cash flow from financing; TR denotes trading revenue; *Trading%* denotes the ratio of trading assets to total assets; and *Total Assets* are in millions. All variables other than R, *Trading%*, and *Total Assets* are deflated by beginning-of-year market value of equity.

[b] All variables, except for *Trading%* and *Total Assets*, are winsorized at three standard deviations from their means.

[c] For bank-years that do not report ΔHFS on the Statement of Cash Flows, ΔHFS is assumed to be zero. The number of non-zero observations for ΔHFS is 189, 21, and 168 for Panels A, B, and C, respectively.

[d] The number of observations for $FCFO3$ is 226, 55, and 171 for Panels A, B, and C, respectively.

[e] The number of observations for TR is 260, 66, and 194 for Panels A, B, and C, respectively.

Spearman = .046) than with $OACC$ (Pearson = −.076; Spearman = −.098). These differential Pearson correlations are driven by the significant positive Pearson correlation between R and ΔTA (.131), because the Pearson correlation between R and ΔHFS is insignificant. In contrast, these differential Spearman correlations are driven by the significant positive Spearman correlation between R and ΔHFS (.138), because the Spearman correlation between R and ΔTA is insignificant. $OCFO$ is positively correlated with both ΔTA (Pearson = .159; Spearman = .161) and ΔHFS (Pearson = .083; Spearman = .088), consistent with banks that generate $OCFO$ investing a portion of that cash flow in net trading assets and HFS loans. CFI is negatively correlated with ΔTA (Pearson = −.109; Spearman = −.022), and CFF is positively correlated with both ΔTA (Pearson = .222; Spearman = .153) and ΔHFS (Pearson = .245; Spearman = .164), consistent with net trading assets displacing other investments because they are financed through the same sources. CFI and CFF are highly negatively correlated (Pearson = −.867; Spearman = −.828), consistent with financing inflows being used to fund investing outflows.[11]

The relatively high correlations among the variables imply that it is necessary to have a fairly large sample of observations in order to generate statistically powerful tests in our regression analyses.[12] Given our small number of banks, this requires pooling observations across a sufficient number of years. Since returns are relatively uncorrelated through time (e.g., the first-order autocorrelation equals .02), this pooling should not yield serially correlated residuals or overstated significance levels in our returns analysis, and in specification analysis we find that this is not the case. We allow for first-order autocorrelation of residuals in our future CFO analysis.

---

[11] These correlations differ from those for nonfinancial firms reported in prior research (e.g., Barth et al. 2001). To illustrate these differences, we estimated the correlations of CFO, ACC, CFI, and CFF from for all the firms on the Compustat industrial tape over our sample period 1991–2003, winsorizing this data at three standard deviations as we do in the paper. We find that the correlation between CFO and ACC is −.24 and the correlation between CFI and CFF is −.53 for this broader sample of firms, i.e., considerably less negative than for our sample. We believe the more negative correlation between CFO and ACC for banks reflect the monetary nature of banks' assets and liabilities, which renders the distinction between cash and accruals less meaningful, since accruals are often very close to cash. We believe the more negative correlation between CFI and CFF for banks reflects their matching of the magnitude of financial assets (investments) and liabilities (financing). On the other hand, we find the correlation between CFO and CFI is −.25 and the correlation between CFO and CFF is −.43 for this broader sample of firms, i.e., considerably more negative than for our banking sample. We believe these more negative correlations reflect nonfinancial firms compensating more strongly for lower CFO by either raising funds externally or reducing investments, and vice versa.

[12] While our variables are highly correlated, none of the variance inflation factors on the variables in our regression equations exceeds 10, the usual threshold for concern about multicollinearity.

TABLE 3
Correlations

Pearson Correlations (upper triangle) and Spearman Correlations (lower triangle)[a, b]

	R	CFO	$FCFO3$	ACC	ΔTA	ΔHFS	$OCFO$	$OACC$	CFI	CFF	TR
R		0.059	**0.182**	0.008	**0.131**	−0.033	**0.176**	−0.076	**−0.159**	0.093	**0.228**
CFO	0.070		**0.158**	**−0.952**	**−0.408**	**−0.725**	**0.418**	**−0.304**	−0.079	**−0.267**	0.019
$FCFO3$	**0.225**	**0.223**		−0.098	0.095	0.026	**0.398**	**−0.376**	**0.172**	**−0.240**	**0.294**
ACC	0.046	**−0.941**	−0.065		**0.411**	**0.796**	**−0.233**	**0.278**	−0.023	**0.354**	0.032
ΔTA	0.042	**−0.368**	0.090	**0.435**		−0.005	**0.159**	**−0.257**	**−0.109**	**0.222**	**0.396**
ΔHFS	**0.138**	**−0.479**	**0.133**	**0.539**	−0.012		0.083	−0.051	0.017	**0.245**	0.002
$OCFO$	**0.262**	**0.467**	**0.525**	**−0.312**	**0.161**	0.088		**−0.843**	**−0.214**	−0.060	**0.478**
$OACC$	−0.098	**−0.424**	**−0.388**	**0.359**	**−0.149**	−0.053	**−0.841**		0.045	0.076	**−0.450**
CFI	**−0.165**	−0.086	0.026	0.021	−0.022	0.035	**−0.159**	0.064		**−0.867**	−0.054
CFF	**0.119**	**−0.249**	−0.080	0.301	**0.153**	**0.164**	−0.026	**0.119**	**−0.828**		0.019
TR	**0.148**	−0.069	0.074	**0.139**	**0.149**	−0.006	**0.139**	**−0.103**	−0.048	0.047	

---

[a] R denotes share return for the 12 months ending on April 30 following the fiscal year-end; CFO denotes cash flow from operations; $FCFO3$ denotes the sum of the next three years' CFO; ACC denotes total accruals, i.e., the difference between net income and CFO; ΔTA denotes the change in net trading assets; ΔHFS denotes the change in held-for-sale loans; $OCFO$ denotes $CFO + \Delta TA + \Delta HFS$; $OACC$ denotes $ACC - \Delta TA - \Delta HFS$; CFI denotes cash flow from investing; CFF denotes cash flow from financing; TR denotes trading revenue. All variables other than R, are deflated by beginning-of-year market value of equity.

[b] Correlations significant at the 5 percent level in a two-tailed test are in boldface.

V. PRIMARY RETURNS AND FUTURE CFO ANALYSIS

Returns Analysis

In this section we estimate and compare the market pricing of the sample banks' cash flow and accrual components. We regress share returns for the 12 months ending on April 30 following the fiscal year on these components aggregated in three different ways. First, as a benchmark and to articulate with prior research by Livnat and Zarowin (1990), we estimate a relatively aggregated model with *CFO*, *ACC*, and the non-operating components *CFI* and *CFF* as explanatory variables. Second, we estimate a disaggregated version of the first model in which the explanatory variables *CFO* and *ACC* are broken into their operating components, *OCFO* and *OACC*, and their hybrid operating and non-operating components, ΔTA and ΔHFS. Third, we estimate a reaggregated version of the second model in which the operating components are replaced with the single explanatory variable $OP = OCFO + OACC$ and the non-operating components are replaced with the single explanatory variable $NONOP = CFI + CFF$.

The third model is attractive because it corresponds in the most direct fashion to our hypotheses *and* because by summing *CFI* and *CFF* it eliminates the strongest source of multi-collinearity among the explanatory variables. For this reason, we tabulate and discuss in detail our hypotheses tests regarding differences between the coefficients on changes in trading assets and HFS loans and the other components of cash flows and accruals only using the third model, although we briefly discuss the untabulated results for the other models as well. Our inferences are the same if we use the second model, however, because as discussed below we find insignificant differences between the coefficients on *OCFO* and *OACC* and between the coefficients on *CFI* and *CFF* in that model.

Specifically, our first returns model, similar to Livnat and Zarowin's (1990) Equation (3), is:[13]

$$R_t = a + b_1 ACC_t + b_2 CFO_t + b_3 CFI_t + b_4 CFF_t + e_t. \qquad (3)$$

In this and subsequent models, the explanatory variables are all deflated by the beginning of fiscal year market value of equity, consistent with prior research using returns models (e.g., Christie 1987). Under Hypothesis H, we expect relatively high positive coefficients on the primarily operating components *ACC* and *CFO* and relatively small positive or negative coefficients on the non-operating components *CFI* and *CFF*. Restating Hypothesis H as restrictions on the coefficients in Equation (3):

H-Equation (3): $b_1 \approx b_2 > b_3 \approx b_4 \approx 0$.

Hypothesis H-Equation (3) is consistent with Livnat and Zarowin's (1990) results for their sample of nonfinancial firms.

Table 4 reports the estimations of the primary returns and future CFO models. As mentioned earlier, all regression variables are winsorized at three standard deviations from

---

[13] Livnat and Zarowin's (1990) explanatory variables in their main regression model are first-differenced. In contrast, our explanatory variables are in levels form, because this appears to yield measures closer to the unexpected portions of those variables for our sample. In particular, the levels of the variables have higher associations with returns both individually and collectively (i.e., in terms of model R^2). We attribute the superior power of a levels model in our setting in part to our decomposition of *CFO* into $OCFO + \Delta TA + \Delta HFS$. If we used a first-differenced model, then the second differences $\Delta(\Delta TA)$ and $\Delta(\Delta HFS)$ would be the explanatory variables. These second differences do not correspond well to the unexpected portions of these variables, presumably due to over-differencing. In any event, our results for Equation (3) are quite consistent with those of Livnat and Zarowin (1990) for their returns model, as discussed below, supporting our specification.

TABLE 4
Regressions of Returns and Future Operating Cash Flows on Cash Flow and Accrual Components[a,b]

Panel A: Estimation

$$R_t = a + b_1 ACC_t + b_2 CFO_t + b_3 CFI_t + b_4 CFF_t + e_t \quad (3)$$

$$R_t = a + b_1 OACC_t + b_2 OCFO_t + b_3 \Delta TA_t + b_4 \Delta HFS_t + b_5 CFI_t + b_6 CFF_t + e_t \quad (4)$$

$$R_t = a + b_1 OP_t + b_2 \Delta TA_t + b_3 \Delta HFS_t + b_4 NONOP_t + e_t \quad (5)$$

$$FCFO3_t = a + b_1 ACC_t + b_2 CFO_t + b_3 CFI_t + b_4 CFF_t + e_t \quad (6)$$

$$FCFO3_t = a + b_1 OACC_t + b_2 OCFO_t + b_3 \Delta TA_t + b_4 \Delta HFS_t + b_5 CFI_t + b_6 CFF_t + e_t \quad (7)$$

$$FCFO3_t = a + b_1 OP_t + b_2 \Delta TA_t + b_3 \Delta HFS_t + b_4 NONOP_t + e_t \quad (8)$$

	Returns			Future Cash Flows[c]		
Model	(3)	(4)	(5)	(6)	(7)	(8)
Intercept	0.175*** (6.4)	0.170*** (6.0)	0.172*** (6.4)	0.015*** (4.2)	0.365*** (3.5)	0.371*** (3.1)
ACC	0.611*** (3.3)			2.324*** (5.1)		
CFO	0.510*** (2.9)			1.864*** (4.4)		
OACC		0.632*** (2.8)			1.229** (2.1)	
OCFO		0.567*** (3.2)			2.037*** (4.7)	
OP			0.616*** (3.7)			1.980*** (4.7)
ΔTA		0.324*** (2.9)	0.321*** (3.2)		0.854*** (4.2)	0.924*** (4.9)
ΔHFS		0.027 (0.3)	0.021 (0.3)		0.388** (2.1)	0.486*** (2.7)
CFI	−0.127** (−2.4)	−0.114** (−2.2)		−0.039 (−0.4)	0.016 (0.2)	
CFF	−0.098* (−1.9)	−0.092* (−1.8)		−0.157* (−1.7)	−0.135 (−1.5)	
NONOP			−0.089* (−1.9)			−0.108 (−1.4)
Adjusted R^2	0.055	0.068	0.070	—	—	—
Likelihood Ratio Test (x^2)	—	—	—	61.33	35.36	57.62
# Observations	330	330	330	226	226	226

Panel B: t-tests of Differences of Coefficients in Equation (5)

	TA	*HFS*	*NONOP*
OP	1.50*	3.00***	4.37***
TA		2.55***	3.34***
HFS			1.00

(*continued on next page*)

TABLE 4 (Continued)

Panel C: t-tests of Differences of Coefficients in Equation (8)

	TA	*HFS*	*NONOP*
OP	2.48***	3.12***	4.92***
TA		1.90**	4.56***
HFS			2.59***

---

[a] *R* denotes share return for the 12 months ending on April 30 following the fiscal year-end; *CFO* denotes cash flow from operations; *FCFO3* denotes the sum of the next three years' *CFO*; *ACC* denotes total accruals, i.e., the difference between net income and *CFO*; ΔTA denotes the change in net trading assets; ΔHFS denotes the change in held-for-sale loans; *OCFO* denotes $CFO + \Delta TA + \Delta HFS$; *OACC* denotes $ACC - \Delta TA - \Delta HFS$; *OP* denotes $OCFO + OACC$; *CFI* denotes cash flow from investing; *CFF* denotes cash flow from financing; *NONOP* denotes $CFI + CFF$. All variables other than *R*, are deflated by beginning-of-year market value of equity.

[b] ***, **, and * indicate statistical significance at 1 percent, 5 percent, and 10 percent, respectively, in a one-tailed test if directional predictions are made and a two-tailed test otherwise.

[c] To mitigate serial dependence of residuals, the estimation of the future cash flow models includes fixed effects for each bank (untabulated) and allows each bank's residuals to follow the same AR(1) process.

their means. In addition, Cook's D test is used to check for any additional outliers in each model. No outlier is identified in the returns regressions, but one outlier is identified in the future CFO regressions and deleted.

The estimation of Equation (3) is reported in the first column of Panel A. Consistent with Hypothesis H-Equation (3), the coefficients on the primarily operating components are significantly and similarly positive, with coefficients of .510 (t = 2.9) on *CFO* and .611 (t = 3.3) on *ACC*, and with these coefficients being insignificantly different from each other.[14] In contrast, but also as hypothesized, the coefficients on the non-operating components are significantly and similarly negative, with coefficients of −.127 (t = −2.4) on *CFI* and −.098 (t = −1.9) on *CFF*, and with these coefficients being insignificantly different from each other. As hypothesized, the coefficients on each of *CFO* and *ACC* are significantly more positive than the coefficients on each of *CFI* and *CFF*. Thus, returns have a relatively strong positive association with the primarily operating components *CFO* and *ACC* and a relatively weak association with the non-operating components *CFI* and *CFF*. These results are qualitatively similar to those of Livnat and Zarowin (1990) for nonfinancial firms.

Our second, disaggregated returns model is:

$$\begin{aligned} R_t = {} & a + b_1 OACC_t + b_2 OCFO_t + b_3 \Delta TA_t \\ & + b_4 \Delta HFS_t + b_5 CFI_t \div b_6 CFF_t + e_t. \end{aligned} \quad (4)$$

Under Hypothesis H, we expect relatively high positive coefficients on the operating components *OACC* and *OCFO*, a lower positive coefficient on ΔTA, an even lower positive coefficient on ΔHFS, and relatively small positive or negative coefficients on the non-operating components *CFI* and *CFF*. Restating Hypothesis H as restrictions on the coefficients in Equation (4):

---

[14] When our hypotheses about the signs of and differences between coefficients are directional, as is usually the case, we say a coefficient or difference of coefficients is significant if the significance level is 5 percent or better in a one-tailed test (i.e., t ≥ 1.65). Otherwise, we evaluate significance at the 5 percent level in a two-tailed test (i.e., t ≥ 1.96).

H-Equation (4): $b_1 \approx b_2 > b_3 > b_4 > b_5 \approx b_6 \approx 0$.

The second column of Table 4, Panel A reports the results of estimating Equation (4). Consistent with Hypothesis H-Equation (4), the coefficients on the operating components are significantly and similarly positive, with a coefficient of .567 (t = 3.2) on *OCFO* and a coefficient of .632 (t = 2.8) on *OACC*, and with these coefficients being insignificantly different from each other. In contrast, but also as hypothesized, the coefficients on the non-operating components are significantly and similarly negative, with coefficients of −.114 (t = −2.2) on *CFI* and −.092 (t = −1.8) on *CFF*, and with these coefficients being insignificantly different from each other. The coefficients on *OCFO* and *OACC* are both significantly more positive than the coefficients on each of *CFI* and *CFF*. The coefficient on the hybrid operating/non-operating component ΔTA is significantly positive at .324 (t = 2.9). As hypothesized, this coefficient is significantly lower than the coefficients on the operating components *OCFO* and *OACC*, and it is significantly higher than the coefficients on the hybrid operating/non-operating component ΔHFS and on the non-operating components *CFI* and *CFF*. The coefficient on the hybrid operating/non-operating component ΔHFS is insignificantly positive at .027. As hypothesized, this coefficient is significantly lower than the coefficients on the operating components *OCFO* and *OACC* and on the hybrid operating/non-operating component ΔTA, but it is insignificantly higher than the coefficients on the non-operating components *CFI* and *CFF*.

Our third, reaggregated returns model is:

$$R_t = a + b_1 OP_t + b_2 \Delta TA_t + b_3 \Delta HFS_t + b_4 NONOP_t + e_t. \qquad (5)$$

Under Hypothesis H, we expect relatively high positive coefficients on *OP*, a lower positive coefficient on ΔTA, an even lower positive coefficient on ΔHFS, and a relatively small positive or negative coefficient on *NONOP*. Restating Hypothesis H as restrictions on the coefficients in Equation (5):

H-Equation (5): $b_1 > b_2 > b_3 > b_4 \approx 0$.

The third column of Table 4, Panel A reports the results of estimating Equation (5). Student t-tests of the differences between the coefficients are reported in Table 4, Panel B. Equation (5) has a higher adjusted R^2 than Equation (4) and so is the best specified of the returns models.

Consistent with Hypothesis H-Equation (5), the coefficient on *OP* is significantly positive at .616 (t = 3.7). The coefficient on *NONOP* is significantly negative at −.089 (t = −1.9). The coefficient on the hybrid operating/non-operating component ΔTA is significantly positive at .321 (t = 3.2). As hypothesized, this coefficient is significantly lower than the coefficient on OP at the 10 percent level (t = 1.5), and it is significantly higher than the coefficients on ΔHFS (t = 2.6) and *NONOP* (t = 3.3). The coefficient on the hybrid operating/non-operating component ΔHFS is insignificantly positive. As hypothesized, this coefficient is significantly lower than the coefficients on *OP* (t = 3.0) and ΔTA (t = 2.6), and it is higher than the coefficient on *NONOP*, though insignificantly so.

In summary, all of the differences of coefficients in Equation (5) have the sign predicted under Hypothesis H, and all the coefficient differences are significant at the 5 percent level except for difference of the coefficients on *OP* versus ΔTA, which is significant at the 10 percent level, and the difference of the coefficients on ΔHFS versus *NONOP*, which is insignificant. Given our maintained assumption that operating activities have stronger valuation implications than non-operating activities, these results imply that the market prices *OP* and ΔTA as significantly more operating than ΔHFS and *NONOP*. The marginally significant coefficient difference for *OP* versus ΔTA suggests that the market prices hybrid trading activities as having a predominantly operating nature, while the insignificant coefficient difference for ΔHFS versus *NONOP* suggests the market price the hybrid activity of originating and selling HFS loans as having a predominantly non-operating nature.

Future CFO Analysis

As in prior research by Wahlen (1994), Dechow et al. (1998), Barth et al. (2001), and Kothari (2001), in this section we attempt to corroborate and explain our prior returns results in terms of the (fundamental) association of the change in net trading assets and other components of cash flows and accruals with future CFO. This analysis should be viewed as complementary with, not subordinate to, the returns analysis; the future CFO analysis indicates whether the components of cash flow and accruals are economically distinct, while the returns analysis indicates whether the market perceives these components to be economically distinct. Moreover, the future CFO analysis is of interest, independent of the returns analysis because SFAC No. 1 (FASB 1978, para. 37) states that a primary goal of financial reporting is to "provide information to help investors, creditors, and others assess the amounts, timing, and uncertainty of prospective net cash inflows to the related enterprise."

We measure future *CFO* in year *t* as the sum of *CFO* for the years $t+1$ to $t+3$ deflated by market value at the beginning of fiscal year *t*, denoted $FCFO3_t$.[15] We measure future *CFO* over a period longer than a year both because returns should reflect the changes in the whole vector of future cash flows and because growth and other factors render the association between the cash flow and accrual components and future *CFO* in any individual year very noisy. As reported in Table 3, the correlation of *FCFO3* with *R* is .182, significant at the 5 percent level.

We regress *FCFO3* on the same three sets of explanatory variables as in the returns regression Equations (3)–(5):

$$FCFO3_t = a + b_1ACC_t + b_2CFO_t + b_3CFI_t + b_4CFF_t + e_t \quad (6)$$

$$FCFO3_t = a + b_1OACC_t + b_2OCFO_t + b_3\Delta TA_t + b_4\Delta HFS_t + b_5CFI_t + b_6CFF_t + e_t \quad (7)$$

$$FCFO3_t = a + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t + e_t. \quad (8)$$

Consistent with the returns regressions, we deflate all the variables in Equations (6)–(8) by beginning of year *t* market value of equity. In untabulated specification analysis, we find that our results are substantively unaffected if we deflate these regressions by average total assets instead, as in Barth et al. (2001). Restating Hypothesis H as restrictions on the coefficients in Equations (6)–(8):

---

[15] To ensure that the time value of money does not affect our results, we conducted a specification test using as the dependent variable the present value of *CFO* in years $t+1$ thorough $t+3$ discounting at a constant 10 percent annual rate. Other than the largely mechanical effect of decreasing the absolute magnitudes of our coefficients, our results are not affected by this specification test.

H-Equation (6): $b_1 \approx b_2 > b_3 > b_4 > b_5 \approx b_6 \approx 0$,

H-Equation (7): $b_1 \approx b_2 > b_3 > b_4 > b_5 \approx b_6 \approx 0$, and

H-Equation (8): $b_1 > b_2 > b_3 > b_4 \approx 0$.

As for the returns analysis, we tabulate and discuss in detail our tests only for the reaggregated model reflected in Hypothesis H-Equation (8), although we briefly discuss the untabulated results for the other models as well.

Because annual *CFO* is serially correlated and *FCFO3* is measured over overlapping three-year periods, the residuals from OLS estimations of Equations (6)–(8) are significantly positively serially correlated; for example, the first-order autocorrelation of the pooled OLS residual in Equation (8) is .65 ($\chi^2 = 58.6$). To address this issue, we estimate these equations using maximum likelihood, allowing the residuals for each firm to follow the same AR(1) process.[16]

The results of all the future *CFO* regressions are reported in Table 4, Panel A, with Equation (6) represented in the fourth column. Consistent with Hypothesis H-Equation (6), the coefficients on *CFO* and *ACC* are significantly positive at 1.864 (t = 4.4) and 2.324 (t = 5.1), respectively, with the two coefficients being insignificantly different. In contrast, but also as hypothesized, the coefficient on the non-operating component *CFI* is insignificantly negative, and the coefficient on the non-operating component *CFF* is significantly negative at the 10 percent level (−.135, t = −1.5); these two coefficients are significantly different from each other in large part due to their high negative correlation. As hypothesized, the coefficients on *CFO* and *ACC* are each significantly more positive than the coefficients on each of *CFI* and *CFF*. Thus, future *CFO* has a relatively strong positive association with the primarily operating components *CFO* and *ACC* and a relatively weak association with the non-operating components *CFI* and *CFF*.

The fifth column of Table 4, Panel A reports the estimation of Equation (7). Consistent with Hypothesis H-Equation (7), the coefficients on the operating components are significantly positive, with a coefficient on *OCFO* of 2.037 (t = 4.7) and a coefficient on *OACC* of 1.229 (t = 2.1), with these two coefficients being significantly different. The coefficients on the non-operating components *CFI* and *CFF* are insignificantly different from zero; these two coefficients are again significantly different from each other. As hypothesized, each of the coefficients on *OCFO* and *OACC* is significantly more positive than the coefficients on each of *CFI* and *CFF*. The coefficient on the hybrid operating/non-operating component ΔTA is significantly positive at .854 (t = 4.2). As hypothesized, this coefficient is significantly lower than the coefficients on *OCFO*, though unexpectedly it is insignificantly lower than the coefficient on *OACC*. As hypothesized, the coefficient on ΔTA is significantly higher than the coefficients on the hybrid operating/non-operating component ΔHFS and the non-operating components *CFI* and *CFI*. The coefficient on the hybrid operating/non-operating component ΔHFS is significantly positive at .388 (t = 2.1). As hypothesized, it is significantly less than the coefficients on the operating components *OCFO* and *OACC* and on the hybrid operating/non-operating component ΔTA, and it is significantly higher than the coefficients on the non-operating components *CFI* (at the 10 percent level) and *CFF*.

---

[16] The actual time-series of the residuals in Equations (6)–(8) and subsequent future CFO models is more complicated than the AR(1) process we assume in the maximum likelihood estimation, because the measurement of *FCFO3* over overlapping three-year periods introduces a MA(2) process into the time-series. Given our relatively small sample size, allowing for such a complicated time-series process for the residuals would yield inefficient estimation of the variance-covariance matrix. For this reason, we assume the simpler AR(1) process, and use the appropriate powers of the estimated first-order autocorrelation parameter of .65 as the higher-order autocorrelations of the residuals. This approach appears to work well. For example, in the OLS estimation of Equation (8), the estimated second-order autocorrelation of the residuals is .47 while the AR(1) process implies this autocorrelation is $.42 = .65^2$, and the estimated third-order autocorrelation is .17 while the AR(1) process implies this autocorrelation is $.27 = .65^3$.

The sixth column of Table 4, Panel A reports the results of estimating Equation (8). Student t-tests of the differences between the coefficients are reported in Table 4, Panel C. Consistent with Hypothesis H-Equation (8), the coefficient on *OP* is significantly positive at 1.980 (t = 4.7) and the coefficient on *NONOP* (−.108) is insignificantly negative. The coefficient on the hybrid operating/non-operating component ΔTA is significantly positive at .924 (t = 4.9). As hypothesized, the coefficient on ΔTA is significantly lower than the coefficient on *OP* (t = 2.5), and significantly higher than the coefficients on ΔHFS (t = 1.9) and *NONOP* (t = 4.6). The coefficient on the hybrid operating/non-operating component ΔHFS is significantly positive at .486 (t = 2.7). As hypothesized, the coefficient on ΔHFS is significantly lower than the coefficients on *OP* (t = 3.1) and ΔTA (t = 1.9), and it is significantly higher than the coefficient on *NONOP* (t = 2.6).

In summary, in the future *CFO* regression Equation (8), the coefficient differences all have the predicted sign and are significant. Specifically, we find that future *CFO* has relatively high positive associations with the operating components of cash flow and accruals, a lower positive association with ΔTA, an even lower positive association with ΔHFS, and insignificant associations with the non-operating components of cash flow.

Summarizing the results of the returns and future *CFO* analyses, all the coefficient differences in both analyses have the correct signs under Hypothesis H, consistent with our argument that trading and originating and selling HFS loans constitute hybrid operating/non-operating activities. Moreover, the results of estimating the returns regression Equation (5) and the future *CFO* regression Equation (8) substantially correspond, consistent with the market being aware, at least to some extent, of the hybrid economic natures of trading activities and HFS loans. The primary disparity between the two analyses is the coefficient differences for *OP* versus ΔTA and for ΔHFS versus *NONOP* are significant in the future *CFO* analysis but are marginally significant or insignificant in the returns analysis. A possible explanation for this disparity is that investors put excess weight on the more predominant aspect of the nature of a hybrid activity, i.e., operating for trading activities and non-operating for HFS loan activities.

VI. SPECIFICATION ANALYSES

Unrealized Gain versus Principal Cash Flows

As discussed above, ΔTA equals the unrealized gain plus the principal cash outflow for trading positions during the period. We hypothesize that the unrealized gain is more operating than the principal cash flow, and so it should have a more positive association with returns and future cash flows. Because we cannot observe either the unrealized gain or the principal cash flows directly,[17] we use trading revenue as proxy for the unrealized gain. Trading revenue is observable for 260 (79 percent) of the 330 sample observations on either

---

[17] Unrealized gains and losses are not observable due to banks' aggregation of unrealized gains and losses with realized gains and losses and other source of non-interest income. The principal cash flows on trading positions are not directly observable due to the classification of these cash flows as operating and banks' universal use of the indirect method of reporting cash flow from operations.

Bank Compustat or the Bank Regulatory Y-9C filings.[18] Trading revenue from these data sources includes realized gains and may include fee income associated with trading; because these additional components of trading revenues are operating, their inclusion does not alter the hypothesis stated above. In addition, banks may differ in the extent to which trading revenue includes fee income. For these reasons, trading revenue is a heterogeneously noisy measure of unrealized gains, likely weakening our results.

Given the relatively small size of our sample, to avoid losing the observations for which trading revenue is not available we set *TR* equal to trading revenue if it is available and to zero otherwise, and add it as an additional explanatory variable in Equations (5) and (8):

$$R_t = a + b_1O_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t + b_5TR_t + e_t \qquad \text{(5-TR)}$$

$$FCFO3_t = a + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t + b_5TR_t + e_t. \qquad \text{(8-TR)}$$

In Equations (5-TR) and (8-TR), the coefficients b_5 on *TR* capture the additional valuation implications of trading revenue, when disclosed, beyond ΔTA, which includes both the unrealized gain and principal cash flow.[19] We hypothesize that *TR* has incremental positive pricing implications beyond ΔTA that are captured in positive coefficients b_5 in Equations (5-TR) and (8-TR):

H-TR: $b_5 > 0$.

Table 5 reports the results of estimating Equations (5-TR) and (8-TR). In Equation (5-TR), the coefficient on *TR* is significantly positive at .723 ($t = 2.0$), consistent with Hypothesis H-TR that the market prices trading revenue as having a more operating nature than the remainder of ΔTA. In Equation (8-TR), the coefficient on *TR* is positive and relatively large (.929 compared to a coefficient of .979 on ΔTA), but not significant. The significant coefficient in *TR* in the returns analysis but insigñificant coefficient on *TR* in the future *CFO* analysis could be due either to the market overvaluing *TR* or to the measurement error in *TR* yielding more attenuation bias in the future *CFO* analysis than in the returns analysis.

Dealer versus Non-Dealer Banks

As discussed above, our sample contains two distinct subsamples: six derivatives dealers with both trading assets and trading liabilities and 31 non-dealer banks with trading assets only. To determine whether these subsamples are heterogeneous, we estimate the reaggregated returns and future *CFO* Equations (5) and (8) allowing the coefficients to

---

[18] The Bank Compustat Annual file (item #120) provides trading revenue for 178 of our observations. Bank Compustat reports trading revenue as it appears in banks' annual financial reports, which likely includes all trading gains and losses and most trading-related fee income, and so it is our preferred source for trading revenue. The bank regulatory Y-9C filings (item #BHCKA220) provide trading revenues for 82 additional observations. The Y-9C filings report trading revenue defined as trading gains and losses plus "incidental income and expense related to the purchase and sale of" trading assets and liabilities; this likely constitutes a subset of fee income streams. For the 89 observations for which trading revenue is observable both on Bank Compustat and in the Y-9C filings, the correlation of the two measures of trading revenue is .80.

[19] We also estimated variants of Equations (5-TR) and (8-TR) with the net principal cash outflow measured as *TR* $- \Delta TA$ in place of ΔTA for the 260-observations sample for which trading revenue is observable. These estimations yield results similar to those reported in Table 5, significantly positive coefficients on *TR* and insignificant coefficients on *TR* $- \Delta TA$.

TABLE 5
Regressions of Returns and Future Cash Flows on Reaggregated Cash Flow and Accrual Components and Trading Revenues[a,b]

$$R_t = a + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t + b_5TR_t + e_t \quad \text{(5-TR)}$$

$$FCFO3_t = a + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t + b_5TR_t + e_t \quad \text{(8-TR)}$$

	Returns	Future Cash Flows[c]
Intercept	0.160*** (5.8)	0.326*** (2.6)
OP	0.556*** (3.3)	2.058*** (4.9)
Δ*TA*	0.237** (2.2)	0.979*** (5.0)
Δ*HFS*	0.004 (0.1)	0.523*** (3.0)
NONOP	−0.081* (−1.7)	−0.153* (−1.8)
TR	0.723** (2.0)	0.929 (.8)
Adj. R^2	0.078	—
Likelihood Ratio Test (x^2)	—	57.08
# Observations	330	226

[a] *R* denotes share return for the 12 months ending on April 30 following the fiscal year-end; *CFO* denotes cash flow from operations; *FCFO3* denotes the sum of the next three years' *CFO*; *ACC* denotes total accruals, i.e., the difference between net income and *CFO*; Δ*TA* denotes the change in net trading assets; Δ*HFS* denotes the change in held-for-sale loans; *OCFO* denotes *CFO* + Δ*TA* + Δ*HFS*; *OACC* denotes *ACC* − Δ*TA* − Δ*HFS*; *OP* denotes *OCFO* + *OACC*; *CFI* denotes cash flow from investing; *CFF* denotes cash flow from financing; *NONOP* denotes *CFI* + *CFF*; *TR* denotes trading revenue. All variables other than *R*, are deflated by beginning-of-year market value of equity.

[b] ***, **, and * indicate statistical significance at 1 percent, 5 percent, and 10 percent, respectively, in a one-tailed test if directional predictions are made and a two-tailed test otherwise.

[c] To mitigate serial dependence of residuals, the estimation of the future cash flow models includes fixed effects for each bank (untabulated) and allows each bank's residuals to follow the same AR(1) process.

differ across the two subsamples.[20] Specifically, we define a dummy variable, *D*, equal to 1 for the six dealers, and 0 otherwise. We interact *D* with the intercept and each of the independent variables in those equations:

$$\begin{aligned} R_t = {} & a + a_DD + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t \\ & + b_5D{*}OP_t + b_6D{*}\Delta TA_t + b_7D{*}\Delta HFS_t + b_8D{*}NONOP_t + e_t \quad \text{(5-D)} \end{aligned}$$

$$\begin{aligned} FCFO3_t = {} & a + a_DD + b_1OP_t + b_2\Delta TA_t + b_3\Delta HFS_t + b_4NONOP_t \\ & + b_5D{*}OP_t + b_6D{*}\Delta TA_t + b_7D{*}\Delta HFS_t + b_8D{*}NONOP_t + e_t. \quad \text{(8-D)} \end{aligned}$$

In these regressions, the coefficients on the uninteracted variables represent the sensitivities for the 31 non-dealer banks; the sums of the coefficients on the uninteracted and interacted variables represent the sensitivities for the six dealers. The coefficients on the interacted variables represent the differences in slopes between the dealers and non-dealer banks.

---

[20] The results of estimating the aggregated Equations (3) and (6) and the disaggregated Equations (4) and (7), distinguishing dealer and non-dealer banks, yield consistent differences across the dealer and non-dealer bank groups as those for the reaggregated Equations (5-D) and (8-D) reported in Table 6.

As a consequence of dealers' focus on fee-generation, we expect the dealers to generate more positive net present value fee streams from their operations than do non-dealer banks. Because most of these fee streams (including those for trading) will be reflected in either *OCFO* or *OACC*, our primary hypothesis is that the coefficient on *OP* is higher for dealers than for non-dealer banks. Because dealers match their trading assets and trading liabilities, thereby diminishing the magnitude of net principal cash flows and thus the non-operating aspects of trading, we also expect a higher coefficient on ΔTA for dealers than non-dealer banks. Finally, loan origination and other fees for HFS loans that must be deferred under SFAS No. 91 (FASB 1986) will be reflected in ΔHFS; if dealer banks have higher fees of this type, then there should also be a higher coefficient on ΔHFS for dealers than for non-dealer banks. Stating these hypotheses as restrictions on the coefficients in Equation (5-D) and (8-D):

H-D-OP: $b_5 > 0$.

H-D-TA: $b_6 > 0$.

H-D-HFS: $b_7 > 0$.

Table 6 reports the results of estimating Equations (5-D) and (8-D). As predicted in Hypothesis H-D-OP, the coefficients on *D*OP* of 1.488 (t = 2.7) in Equation (5-D) and 2.487 (t = 2.0) in Equation (8-D) are both significantly positive, consistent with *OP* having greater valuation implications for the dealers than for the non-dealer banks. As predicted in Hypothesis H-D-HFS, the coefficients on $D^*\Delta HFS$ of .783 (t = 2.6) in Equation (5-D) and 1.316 (t = 2.3) in Equation (8-D) are both significantly positive, consistent with ΔHFS having greater valuation implications for the dealers than for the non-dealer banks. In contrast, the coefficient on the interacted variable $D^*\Delta TA$ is insignificant in Equation (5-D) and significantly positive at the 10 percent level (.611; t = 1.3) in Equation (8-D), only partially and weakly consistent with Hypothesis H-D-TA that trading has a more operating character for dealers than for non-dealer banks.

The consistency of the results across the returns and future cash flow equations in this section again suggests that the market has a general understanding of the different valuation implications of the various cash flow and accrual components.

Other Specification Analyses

We conducted the following additional specification analyses to ensure that our results are robust. First, banks merged or acquired each other continuously throughout the sample period. Pooling mergers yield restatements of historical financial statements, which along with restatements for any other reason constitute an issue when we collect prior years' data from a given year's financial report filed on EDGAR, as discussed in Section IV. To eliminate significant mergers and acquisitions and restatements, we deleted bank-year observations for which assets grew more than 20 percent during the year or for which annual earnings from Bank Compustat (originally reported) differs from the annual earnings from EDGAR (potentially restated). This eliminated 20 percent of the observations in the primary returns analyses. The results are similar for all analyses except for those involving *TR* reported in Table 5, for which the coefficient on *TR* becomes insignificant.

Second, because the observations arise from a single industry, they are cross-correlated in a given year. In our view, the common industry effects have the same interest as the

TABLE 6
Regressions of Returns and Future Cash Flows on Reaggregated Cash Flow and Accrual Components Distinguishing Dealer and Non-Dealer Banks[a,b]

$$R_t = a + a_D D + b_1 OP_t + b_2 \Delta TA_t + b_3 \Delta HFS_t + b_4 NONOP_t + b_5 D^*OP_t + b_6 D^*\Delta TA_t + b_7 D^*\Delta HFS_t + b_8 D^*NONOP_t + e_t \qquad \text{(5-D)}$$

$$FCFO3_t = a + a_D D + b_1 OP_t + b_2 \Delta TA_t + b_3 \Delta HFS_t + b_4 NONOP_t + b_5 D^*OP_t + b_6 D^*\Delta TA_t + b_7 D^*\Delta HFS_t + b_8 D^*NONOP_t + e_t \qquad \text{(8-D)}$$

	Returns	**Future Cash Flows**[c]
Intercept	0.192***	0.363***
	(6.7)	(2.8)
OP	0.502***	1.068**
	(2.8)	(2.0)
D	−0.227***	0.132
	(−2.7)	(0.4)
ΔTA	0.319***	0.273
	(2.4)	(0.9)
ΔHFS	−0.023	0.359**
	(−0.3)	(1.9)
NONOP	−0.074	−0.095
	(−1.5)	(−1.1)
*D**OP*	1.488***	2.487**
	(2.7)	(2.0)
$D^*\Delta TA$	−0.097	0.611*
	(−0.4)	(1.3)
$D^*\Delta HFS$	0.783**	1.316**
	(2.5)	(2.3)
*D**NONOP*	−0.013	0.066
	(−0.1)	(0.3)
Adjusted R^2	0.096	—
Likelihood Ratio Test (χ^2)	—	50.23
# Observations	330	226

---

[a] *R* denotes share return for the 12 months ending on April 30 following the fiscal year-end; *CFO* denotes cash flow from operations; *FCFO3* denotes the sum of the next three years' *CFO*; *ACC* denotes total accruals, i.e., the difference between net income and *CFO*; ΔTA denotes the change in net trading assets; ΔHFS denotes the change in held-for-sale loans; *OCFO* denotes $CFO + \Delta TA + \Delta HFS$; *OACC* denotes $ACC - \Delta TA - \Delta HFS$; *OP* denotes $OCFO + OACC$; *CFI* denotes cash flow from investing; *CFF* denotes cash flow from financing; *NONOP* denotes $CFI + CFF$; *D* takes a value of 1 for dealer banks, and 0 for non-dealer banks. All variables other than *R* and *D* are deflated by beginning-of-year market value of equity.

[b] ***, **, and * indicate statistical significance at 1 percent, 5 percent, and 10 percent, respectively, in a one-tailed test if directional predictions are made and a two-tailed test otherwise.

[c] To mitigate serial dependence of residuals, the estimation of the future cash flow models includes fixed effects for each bank (untabulated) and allows each bank's residuals to follow the same AR(1) process.

firm-specific effects, so that removing those effects eliminates variation of interest. However, we also estimated all models using fixed time effects and using a robust variance estimation approach that allows for correlation of residuals across clusters of observations, in this case the observations in a given year. The estimated coefficients have the hypothesized relationships in these analyses, although the t-statistics generally diminish by about a third.

VII. CONCLUSION

In this paper we argue that banks' trading activities have a hybrid operating/non-operating nature that is not fully captured by the classification of the cash flows on these activities as operating during our sample period from 1991–2003 under SFAS No. 102 and related accounting practices. Reflecting this hybrid nature, we hypothesize that the change in net trading assets has a less positive association with returns and future CFO than do the pure operating components of cash flows and accruals and a more positive association with returns and CFO than do the pure non-operating components of cash flows. We find support for this hypothesis on a sample of U.S. banks holding an appreciable amount of trading assets.

We conduct three additional empirical analyses to further investigate how the market prices hybrid operating/non-operating flows and how these flows relate to future CFO. First, we hypothesize and find that the market prices the change in net trading assets as more operating than the change in held-for-sale loans, another important hybrid but more investing-oriented activity. Second, for the observations for which we can observe the operating and non-operating components of the change in net trading assets, we hypothesize and find that returns and future cash flows are more positively associated with the trading revenue (operating) component than with the principal cash flow (non-operating) component. Third, to address potential heterogeneity in our sample and to examine whether the market differentiates firms based on the relative operating nature of their activities, we hypothesize and find that derivative dealers' activities generally are more operating than are those of non-dealer banks, although the results for the change in net trading assets are only partially and weakly consistent with our hypotheses.

Collectively, our market pricing and future cash flow results suggest that the market appreciates the hybrid operating/non-operating nature of trading activities and also of HFS loans and differentiates firms based on the relative operating nature of their activities. This appreciation likely reflects the fact that investors can observe the changes in trading positions and HFS loans separately from other CFO and accruals, and so they can make adjustments for these changes in their valuation analyses. Our results imply that it is important that the cash flows and accruals on hybrid activities, such as trading positions and HFS loans, be reported in a disaggregated fashion on the SCF or elsewhere in financial reports, regardless of whether these items are classified as operating.

However, our results also imply that any single classification of the cash flows on trading positions and HFS loans cannot logically capture the hybrid economic natures of these activities. Hence, additional disclosures regarding the natures and financial statement effects of these and other hybrid activities would be useful to investors and other users of financial reports. Specifically, it would be useful to require disclosure of the amounts of principal cash flows and timing differences (unrealized gains and losses) associated with these activities; these amounts generally cannot be inferred from current disclosures of the net changes in these accounts.

Our results contribute to the literature on the market pricing of the components of cash flow and accruals, which until now has primarily focused on nonfinancial firms. Our paper is the first to hypothesize and document empirically that the cash flow from operations of financial institutions and trading firms include components with hybrid operating/non-operating natures that are priced by the market as such. While we examine a relatively small sample of banks with significant trading operations, our results have more general importance because a broad set of financial institutions and trading firms have cash flows whose hybrid nature is not fully captured by their SCF classification. For example, lessors classify the cash flows on operating leases and insurers classify the cash flows on traditional insurance contracts as operating, despite the partly non-operating natures of these activities. Financial institutions and trading firms play a large and increasing role in the overall economy, and so the SCF classification issues we examine are likely to become more important over time.

Our results are of current interest given the recent highly publicized cases in which Enron and other firms issued prepay liabilities that raised funds that were classified as operating cash inflows under prior accounting practices. To eliminate firms' ability to use prepays to increase cash flow from operations, FASB required in SFAS No. 149 that the cash flows on derivative liabilities with other-than-insignificant initial value be classified as financing, and it is likely that this requirement will be applied to trading liabilities generally. While FASB's motivation for SFAS No. 149's requirement is clear and understandable, this requirement is conceptually problematic, especially for trading firms with matched trading assets and liabilities, because it will cause the cash flows on these trading assets and trading liabilities to be classified inconsistently on the SCF. In particular, growing trading firms will have operating cash outflows for trading assets and financing cash inflows for trading liabilities.

We believe that FASB should consider remedying this inconsistency by requiring the cash flows on matched trading assets and liabilities to be classified in a consistent fashion on the SCF. There are two ways FASB could do this: (1) it could require the cash flows on matched trading assets and liabilities to be classified as operating (i.e., partly or wholly undo the requirement of SFAS No. 149 and return to prior accounting practices) or (2) it could require the (principal) cash flows on trading assets to be classified as investing (i.e., modify the requirements of SFAS No. 102) and on trading liabilities to be classified as financing (analogous to SFAS No. 149). As discussed above, our results imply that the market views the cash flows on trading positions as largely though not entirely operating, and so on this basis the first approach appears to be more economically descriptive. On the other hand, the second approach might be more robust to cases in which trading assets and liabilities are not matched.

APPENDIX

Statement of Cash Flows
M&T Bank Corporation for Fiscal Year 2000
(in thousands)

Cash flows from operating activities:		
Net income (NI)	$ 286,156	
+ Provision for credit losses		38,000
+ Depreciation and amortization of premises and equipment		30,164
+ Amortization of capitalized servicing rights		24,392
+ Amortization of goodwill and core deposit intangible		69,576
+ Provision for deferred income taxes		(5,911)
+ Asset write-downs		1,674
− Net gain on sales of assets		(6,631)
− Net change in accrued interest receivable (payable)		25,540
− Net change in other accrued income (expense)		(27,901)
− Ordinary accruals (*OACC*)	148,903	
= Ordinary cash flow from operations (*OCFO*)[a]	435,059	
− Net change in trading assets and liabilities (ΔTA)[a]	(6,868)	
− Net change in loans held for sale (ΔHFS)[a]	(81,549)	
= Cash flow from operations (*CFO*)	346,642	

Cash flows from investing activities:	
Proceeds from sales of investment securities	771,815
Proceeds from maturities of investment securities	498,125
Purchase of investment securities	(458,421)
Additions to capitalized servicing rights	(33,694)
Net increase in loans and leases	(1,467,187)
Capital expenditures, net	(18,784)
Acquisitions, net of cash acquired	169,912
Purchase of bank owned life insurance	(35,000)
Other	13,183
= Cash flow from Investing Activities (*CFI*)	(560,051)
Cash flows from financing activities:	
Net decrease in deposits	(321,735)
Net increase (decrease) in short-term borrowings	(830,214)
Proceeds from long-term borrowings	1,000,896
Payments on long-term borrowings	(32,224)
Purchases of treasury stock	(54,947)
Dividends paid—common	(51,987)
Other	34,830
= Cash flow from financing activities (*CFF*)	(255,381)

---

[a] Accruals (ACC) = $OACC + \Delta TA + \Delta HFS$ = –\$148,903 + \$6,868 + \$81,549 = –\$60,468.

REFERENCES

Ali, A. 1994. The incremental information content of earnings, working capital from operations, and cash flows. *Journal of Accounting Research* 32: 61–74.

American Institute of Certified Public Accountants (AICPA). 1998. Audit and Accounting Guide: Banks and Savings Institutions, with conforming changes as of May 1, 1998. New York, NY: American Institute of Certified Public Accountants.

American Institute of Certified Public Accountants (AICPA). 1998. Audit and Accounting Guide: Brokers and Dealers in Securities, with conforming changes as of May 1, 1998. New York, NY: American Institute of Certified Public Accountants.

Barth, M., W. Beaver, and M. Wolfson. 1990. Components of earnings and the structure of bank share prices. *Financial Analysts Journal* (May-June): 53–60.

———. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69: 1–25.

———, D. Cram, and K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76: 27–58.

Beaver, W., and M. McNichols. 2001. Do stock prices of property casualty insurers fully reflect information about earnings, accruals, cash flows, and development? *Review of Accounting Studies* 6: 197–220.

Bernard, V., and T. Stober. 1989. The nature and amount of information reflected in cash flows and accruals. *The Accounting Review* 64: 624–652.

Bowen, R., D. Burghstahler, and L. Daley. 1987. The incremental information content of accruals versus cash flows. *The Accounting Review* 62: 723–747.

Carlson, M., and R. Perli. 2004. Profits and balance sheet developments at U.S. commercial banks in 2003. *Federal Reserve Bulletin* (Spring): 162–191.

Cheng, Q., P. Frischmann, and T. Warfield. 2000. The market perception of corporate claims. Working paper, University of Wisconsin.

Christie, A. 1987. On cross-sectional analysis in accounting research. *Journal of Accounting and Economics* 9: 231–258.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18: 3–42.

———, S. P. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

Financial Accounting Standards Board (FASB). 1978. *Objectives of Financial Reporting by Business Enterprises*. Statement of Financial Accounting Concepts No. 1. Norwalk CT: FASB.

———. 1986. *Accounting for Nonrefundable Fees and Costs Associated with Originating or Acquiring Loans and Initial Direct Costs of Leases*. Statement of Financial Accounting Standards No. 91. Norwalk CT: FASB.

———. 1987. *Statement of Cash Flows*. Statement of Financial Accounting Standards No. 95. Norwalk CT: FASB.

———. 1989. *Statement of Cash Flows—Exemption of Certain Enterprises and Classification of Cash Flows from Certain Securities Acquired for Resale*. Statement of Financial Accounting Standards No. 102. Norwalk CT: FASB.

———. 1993. *Accounting for Certain Investments in Debt and Equity Securities*. Statement of Financial Accounting Standards No. 115. Norwalk CT: FASB.

———. 1994. *Disclosure about Derivative Financial Instruments and Fair Value of Financial Instruments*. Statement of Financial Accounting Standards No. 119. Norwalk CT: FASB.

———. 2003. *Amendment of Statement 133 on Derivative Instruments and Hedging Activities*. Statement of Financial Accounting Standards No. 149. Norwalk CT: FASB.

Foster, G. 1977. Valuation parameters of property-liability companies. *The Journal of Finance* 32: 823–835.

G4+1. 1999. Reporting financial performance. Discussion paper. August. London, U.K.: International Accounting Standards Committee.

Hopkins, P. 1996. The effect of financial statement classification of hybrid financial instruments on financial analysts' stock price judgments. *Journal of Accounting Research* 19: 33–50.

Jorion, P. 2002. How informative are value-at-risk disclosures? *The Accounting Review* 77: 911–931.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31: 105–231.

Linsmeier, T., C. Shakespeare, and T. Sougiannis. 2000. Liability/equity classifications and shareholder valuation. Working paper, Michigan State University.

Liu, C., S. Ryan, and H. Tan. 2004. How banks' value-at-risk disclosures predict their total and priced risk: Effects of bank technical sophistication and learning over time. *Review of Accounting Studies* 9: 265–294.

Livnat, J., and P. Zarowin. 1990. The incremental information content of cash flow components. *Journal of Accounting and Economics* 13: 25–46.

Nurnberg, H. 1993. Inconsistencies and ambiguities in cash flow statements under FASB Statement No. 95. *Accounting Horizons* 7: 60–75.

Office of Thrift Supervision. 2004. *2003 Fact Book: A Statistical Profile of the Thrift Industry*. Washington, DC: Office of Thrift Supervision.

Rayburn, J. 1986. The association of operating cash flow and accruals with security returns. *Journal of Accounting Research* 24: 112–138.

Ryan, S. 2002. *Financial Instruments and Institutions: Accounting and Disclosure Rules*. Hoboken, NJ: John Wiley & Sons.

Wahlen, J. 1994. The nature and information in commercial bank loan loss disclosures. *The Accounting Review* 69: 455–478.

Wilson, G. P. 1986. The relative information content of accruals and cash flows: Combined evidence at the earnings announcement and annual report release date. *Journal of Accounting Research* 24: 165–200.

———. 1987. The incremental information content of accrual and funds components of earnings after controlling for earnings. *The Accounting Review* 62: 293–322.

THE ACCOUNTING REVIEW
Vol. 81, No. 2
2006
pp. 473–495

Determinants of Audit Preparers' Workpaper Justifications

Premila Gowri Shankar
Nanyang Technological University

Hun-Tong Tan
Nanyang Technological University

ABSTRACT: Research indicates that justifiability of a decision is a key determinant of work quality. In this paper, we investigate whether and how the requirement for preparers to justify an audit conclusion to a reviewer with task preferences that are similar to or dissimilar from their explicitly stated initial task preferences determine the nature and extent of justifications documented in their workpapers, and how this effect is moderated by their technical and tacit managerial knowledge. Results of an experiment with auditors as participants indicate that auditors use different forms of justifications under different circumstances. Auditors with high tacit managerial knowledge and high technical knowledge have a higher proclivity to enumerate more pro versus con reasons, and to consider a wider breadth of issues when they are required to justify their conclusions to reviewers with *dissimilar* task preferences. In contrast, high tacit managerial knowledge auditors required to justify to reviewers with *similar* initial task preferences have a greater tendency to employ an evidence framing approach, in which language is used to emphasize consistent evidence by downplaying associated inconsistent evidence. In additional analyses, we find that these three specific forms of justifications are positively associated with evaluations of work quality.

Keywords: *justification; tacit managerial knowledge; review process.*

Data Availability: *Contact the authors.*

I. INTRODUCTION

The need to justify a conclusion or point of view is a ubiquitous feature in many business and professional contexts. In such settings, the justifier generally seeks to provide support for an expressed viewpoint, and to persuade recipients of its validity. For example, managers routinely justify their assessments of the company's past and future outlook, and may adopt persuasive strategies in justifying their assessments to investors and analysts (Sedor 2002; Kuperman 2003). Financial analysts also provide justifications for their research reports, and investors' reliance on these reports is determined in part by the persuasiveness of these justifications (Hirst et al. 1995).

We appreciate the helpful comments of Amit Das, Jun Han, Karim Jamal, Terence Ng, Pearl Tan, Ken Trotman, workshop participants at Nanyang Technological University, Jane Kennedy (editor), and two anonymous reviewers. We thank the auditors from the two Big 4 public accounting firms for participating in our study.

Editor's note: This paper was accepted by S. Jane Kennedy, Editor.

Submitted June 2004
Accepted August 2005

In this paper, we focus on justifications in an auditing context, where the quality of documented justifications is especially important. Auditing standards mandate that workpapers contain sufficient and appropriate justifications that support a documented conclusion (SAS No. 96, AICPA 2002). Furthermore, such justifications are referred to as a source of evidence by others such as judges, jurors, expert witnesses, and regulators in litigation cases (Kennedy et al. 1997; Beus 2003; Ricchiute 2004). In particular, for ambiguous unstructured audit tasks with no one single appropriate solution, justifiability or defensibility of the documented judgment or decision is a major measure of work quality (Waller and Felix 1984; Emby and Gibbins 1988; Kennedy et al. 1997).

Despite its importance, few studies examine the nature and determinants of justifications documented by workpaper preparers. Two studies are relevant. Koonce et al. (1995) and Tan et al. (1997) provide evidence suggesting that preparers increase the quantity of justifications when they are made accountable to a reviewer with unknown preferences. However, both papers document effects on the *quantity* not the *nature* of preparers' justifications, and preparers have been posited to vary or stylize the *nature* of their workpaper documentation as part of their reputation management strategies (Rich et al. 1997; Gibbins and Trotman 2002).[1]

Current trends toward real-time reviews involve greater and earlier interactions among reviewers and preparers, where reviewers provide front-end guidance to preparers in the conduct of the audit (Rich et al. 1997; Gibbins and Trotman 2002). One concern arising from real-time reviews is that reviewers communicate their task preferences (that is, their preferred audit conclusions on the task) to workpaper preparers *too* early in the audit process (Rich et al. 1997; Wilks 2002). Prior studies suggest a cost to real-time reviews in that preparers required to justify to a reviewer with a particular directional preference are systematically influenced by knowledge of their reviewer's preference in terms of their hypothesis generation (Peecher 1996) and evidence interpretation (Wilks 2002). Note that preparers in these studies do not form initial audit judgments or conclusions related to the task before knowing the reviewer's preference.

In a departure from these prior studies, we investigate conditions in which preparers do explicitly state their task preferences (that is, their preferred audit conclusions on the task) before knowing the reviewers' task preferences. Specifically, our first factor—justification requirement—refers to an accountability-inducing requirement (Tetlock 1985a, 1992, 2002) where the preparer is required to justify his/her audit conclusion to a reviewer whose task preference is either similar to or dissimilar from the explicitly stated initial task preference of the preparer. Our motivation for looking at this factor can be viewed from two perspectives. First, an inference from prior research is that the costs associated with real-time review arise in part because preparers now have fewer opportunities to form opinions that are dissimilar from that made by the reviewer (in that preparers form a task preference only after knowing the reviewer's task preference). From this perspective, it is important to know the effects when preparers are *required* to explicitly form their task preferences *prior to* the reviewer's task preference being communicated to them. Such a requirement naturally increases the likelihood that the preparers form task preferences that are dissimilar from that formed by the reviewer, and findings on the efficacy of this requirement would be useful to CPA firms considering firm policies or interventions that mitigate limitations of real-time reviews. Second, in practice and particularly during periods when reviewers face the practical constraints of tight reporting deadlines and involvement in several concurrent audit assignments, reviewers likely communicate their conclusions or opinions only *after* preparers have collected evidence and documented their initial conclusions.[2] Thus, there are circumstances when preparers naturally have to form their task preferences before knowing the reviewers' task preferences. For these reasons, we expand on the dimension of justification requirement, and consider the match or mismatch between the initial task preference of the reviewer and that explicitly stated by the preparer; that is, whether the justification requirement is to a reviewer with a task preference that is similar to or dissimilar from that of the preparer. Because preparers' responses to these justification requirements are likely constrained by the skill sets they possess, we investigate the moderating effects of two personal attributes: technical knowledge and tacit managerial knowledge (specifically, knowledge on managing their superior's impressions).

---

[1] Koonce et al. (1995) find no justification effects on audit procedures and processes (e.g., problem identification, information request) adopted by preparers.

We conducted an experiment in which audit preparers performed an ambiguous audit task with no one single appropriate solution (doubtful accounts task). In the experiment, preparers assessed the appropriateness of the client's accounting treatment for the allowance for doubtful accounts, and explicitly stated their initial task preferences. They were then informed of their superior's task preferences, and wrote memos to their superior justifying their final decisions.

Results show that preparers attempt to justify their conclusions in one of three ways: (1) presentation of a greater number of arguments in support of versus against their documented conclusion (*net persuasive evidence*); (2) consideration of a broad spectrum of perspectives (e.g., client-risk, internal control risk, GAAP, materiality, etc) (*breadth of issues*); and (3) using language to frame evidence so as to highlight evidence consistent with the documented conclusion and downplay evidence inconsistent with the documented conclusion (*evidence framing*). We find that preparers are strategic in their use of these different forms of justifications, depending on the context and the personal attributes of the preparers. For the first two justification forms, technical and tacit managerial knowledge moderate the effects of justification requirements. Preparers document more (less) net persuasive evidence when they are required to justify to a reviewer with dissimilar (similar) task preferences, but only when preparers possess high tacit managerial and high technical knowledge. We find this same pattern of results with respect to the documentation of breadth of issues; in addition, we find that preparers with high tacit managerial but low technical knowledge also document greater breadth of issues when they are required to justify to a reviewer with similar task preferences. The first two forms (net persuasive evidence and breadth of issues) involve the listing of relevant evidence related to the auditing issue at hand. In contrast, the third form, evidence framing, relates to the use of language to frame pro and con evidence. We find that the determinants of and directional effects on evidence framing are different from the other two justification forms. In contrast to the other two justification forms, evidence framing is used to a greater extent only when high tacit managerial knowledge preparers are required to justify to a reviewer with *similar* (as opposed to dissimilar) task preferences. Further analysis suggests that these preparers appear to economize on the enumeration of pro versus con reasons by treating evidence framing as a substitute for net persuasive evidence when justifying to a similar-preference reviewer. We also provide evidence that all three forms of justifications have direct and positive effects on evaluators' assessments of work quality.

---

[2] These observations were gathered from interviews with audit partners and managers from the Big 4 firms. They also noted that depending on the audit task and individual auditor, there would be variations in terms of when the reviewer's conclusion is communicated to the preparer (i.e., before or after the preparer has formed an initial conclusion).

Our study contributes to the accounting and psychology literature in a number of ways. Prior research in accounting (e.g., Peecher 1996; Wilks 2002) focuses on the construct, justification to a superior with known preferences, and shows that it systematically affects preparers' cognition or judgment. We expand on this construct, and demonstrate that what matters to the preparer is not simply the reviewer's task preference per se, but the match or mismatch between the reviewer's task preference and the initial task preference explicitly stated by the preparer. A similar point can be made to the related prior psychology literature on bolstering and escalating commitment. This literature documents that when people have explicitly stated their initial decision, they have a heightened tendency to escalate commitment to and justify the correctness of that initial decision, even in the face of negative feedback (Staw 1976; Fox and Staw 1979; Caldwell and O'Reilly 1982; Brockner 1992). Our results suggest that the extent of bolstering or escalating commitment to an explicitly stated initial decision may depend on the match or mismatch between this explicitly stated initial decision and the preferences of the justifiee.[3]

Prior research also suggests that knowledge of the reviewers' task preferences leads to dysfunctional effects on preparers' performance (Peecher 1996; Wilks 2002). Our study suggests that dysfunctional effects may not necessarily be the case. We provide evidence that dysfunctional effects may occur if the reviewers' communicated task preferences are similar to that of the preparers, *but not* when these preferences are dissimilar. In fact, our results suggest that benefits may actually accrue in the latter case with respect to listing "hard" evidence related to pro versus con reasons and auditing/accounting issues.

In addition, while the separate effects on auditors' performance of technical knowledge and tacit managerial knowledge have been examined (Tan and Libby 1997), there is no empirical evidence on the interactive effects of these two variables, nor how they jointly interact with justification requirements.[4] The closest study we are aware of is by Tan and Kao (1999), who show that auditors' accountability to a superior with *unknown* preferences interacts with technical knowledge. They do not examine the effect of tacit managerial knowledge. Our results imply that the effects of accountability and technical knowledge are more complex than what has been documented in the extant literature, and that these findings may be relatively incomplete without considering the moderating role of tacit managerial knowledge.

Finally, relatively little research has been done on the determinants of justification memo quality. A recent tax study finds that various aspects of tax knowledge are associated with adaptive information search, which in turn is associated with the quality of justification memos (Magro 2005). However, her study does not examine the specific forms of justifications that may be adopted by tax professionals in documenting research memos, nor does it examine the interactive effects of technical knowledge with tacit managerial knowledge and justification requirements. Our study extends the literature on justifications by documenting the kinds of justifications that preparers naturally engage in, and by providing evidence that the nature and extent of these justifications is a complex function of the attributes of the preparers and justification requirements. We also provide evidence that justification memos using different justification strategies influence memo quality, a finding that has important implications on auditors' defense of their work quality.

---

[3] Much of the escalating commitment literature invariably focuses on dependent variables such as resource allocation assessments (see review by Brockner 1992), rather than justification strategies or justification quality related to technical issues.

[4] Tan and Libby (1997) focus on how auditors of different performance ranks (e.g., highly ranked manager versus mediocre manager) have different levels of technical knowledge or tacit managerial knowledge. They do not examine how these two variables interactively determine performance.

The rest of the paper is organized as follows. We develop our hypotheses in Section II, describe the research method in Section III, report the results in Section IV, and conclude with a discussion of the implications and limitations of the paper in Section V.

II. BACKGROUND AND HYPOTHESES DEVELOPMENT

Justification Forms

In formulating a justification, the justifier attempts to provide support for an expressed viewpoint. Put another way, the justifier wants to persuade the target audience that his/her viewpoint is valid. While justifications in auditing settings have been previously investigated (Koonce et al. 1995; Tan et al. 1997), there is little empirical evidence on the forms of justifications employed by workpaper preparers, or features that make a justification persuasive in the eyes of the reviewer.

Justifications can be viewed as a form of reasoned argument, such that a justification is more persuasive the more "reasoned" or logical its argument quality (Petty et al. 1983; Petty and Wegener 1991). For instance, following the jurisprudence model by Toulmin and colleagues (Toulmin 1958; Toulmin et al. 1984), a reasoned argument consists of a claim, along with (a) grounds, considerations, and backings; and (b) rebuttals that acknowledge the presence of certain facts/evidence that may invalidate or weaken the claim.[5] Toulmin et al.'s (1984) work suggests that an argument or justification is persuasive if the logical components identified above are present. Conditions (a) and (b) together suggest that listing *both* reasons supporting (pro) and refuting (con) the claim can be viewed as providing reasoned argumentation. Indeed, the persuasion literature demonstrates that two-sided arguments (involving both pro and con arguments) are more persuasive than one-sided arguments (incorporating only pro arguments) in that the former approach reduces the likelihood of counter-argumentation and enhances message credibility (Pechmann 1992). However, if too many con arguments are included, then the persuasiveness of the message may be eroded (Toulmin et al. 1984; Golden and Alpert 1987). Hence, we expect reasoned arguments to have more pro than con evidence (net persuasive evidence).

Justifications that carefully assess all considerations (pro and con reasons) related to the claim (as suggested by Toulmin et al. [1984]) also likely involve an assessment of a broad variety of issues from different angles. In an auditing context, auditors are trained to arrive at their conclusions after considering the implications of a variety of issues such as the client's business risk, industry risk, businesses processes and controls, materiality, and GAAP (e.g., see Winograd et al. 2000). All things considered, the basis for the conclusion can be seen to be comprehensive and persuasive the more such issues are considered; we refer to this as breadth of issues.

Other than the listing of evidence (e.g., net persuasive evidence and breadth of issues), argument presentation by way of language influences its persuasiveness (Hirschberg 1990). More recently, Areni (2003) draws upon the literature in logic, social psychology, jurisprudence, and sociolinguistics to suggest that persuasive arguments may be viewed as those that have certain grammatical features in addition to structural features identified by Toulmin et al. (1984). Specifically, justifications may be framed by using language to downplay evidence that is inconsistent with one's conclusion by countering it with related pieces of consistent evidence (Kamins and Assael 1987; Hirschberg 1990; Crowley and Hoyer 1994; see also Rich et al. 1997). This form of framing has been found to influence auditors' judgments (Tan and Yip-Ow 2001). For example, consider the following sentence justifying the need for additional allowance for doubtful accounts: "notwithstanding the use of experienced credit agents, companies with lower credit ratings are more likely to go bust." A few features of this justification should be noted. First, there is strategic placement of the evidence. The merit of the evidence related to the use of experienced credit agents (inconsistent with the documented conclusion) is downplayed by juxtaposing it with a piece of evidence that is consistent with the desired conclusion (greater allowance for doubtful accounts because accounts receivable may become uncollectible). Second, there is the use of semantics in that the word "notwithstanding" serves to link the inconsistent and consistent evidence, and to downplay the merits of the inconsistent evidence.

---

[5] According to Toulmin et al. (1984), claims can have associated qualifiers (e.g., use of words like possibly, certainly, etc.) to reflect the strength of the claim. For example, in the case of some accounting issues such as contingent liabilities, the use of probability phrases is prescribed by authoritative guidance.

Overall, we examine three forms of justifications that may be used by preparers to enhance the persuasiveness of their justifications: net persuasive evidence, breadth of issues, and evidence framing.[6] We discuss determinants of their usage below.

Determinants of Justification Form Usage

An extensive body of literature shows that accountability, the need to justify decisions to an audience, generally leads to greater effort and vigilance (Tetlock 1985a, 1992, 2002). Prior accounting studies show that accountability to a superior with unknown preferences has debiasing properties and leads to improved performance (Kennedy 1993; Tan and Kao 1999). Research also shows that the need to justify decisions to a superior with known preferences leads auditors to align their judgments and interpret evidence in a manner consistent with the superior's preferences (Peecher 1996; Turner 2001; Wilks 2002). However, as we posit below, the effects of justification are likely dependent on the match or mismatch between the preparer's explicitly stated initial preference and the preference of the reviewer.

The literature on escalating commitment documents a heightened tendency for individuals to justify the correctness of their explicitly stated initial decisions in the face of negative feedback, particularly when they feel personally responsible for the initial decision (Staw 1976; Fox and Staw 1979; Caldwell and O'Reilly 1982; Brockner 1992). Social psychology research also shows that individuals are more (less) likely to engage in deep cognitive processing when they are exposed to information that is inconsistent (as opposed to consistent) with their initial beliefs (Pyszczynski and Greenberg 1981; Wong and Weiner 1981; Hastie 1984). Individuals exposed to information that is inconsistent with their initial beliefs engage in greater cognitive processing of information to either "explain away" the information or to change their beliefs (Pyszczynski and Greenberg 1981; Wong and Weiner 1981; Hastie 1984). This cognitive effect is likely stronger when the challenge to one's initial belief comes from a superior or evaluator (e.g., a workpaper preparer's reviewer; Tetlock 1985a, 1985b). Moreover, in the auditing context, because the preparer's justification requirement is to a superior who evaluates his/her documented work, we anticipate greater effort to be exerted in the preparation of the justification memo.

---

[6] Rich et al. (1997) also identify other strategies such as selective failure to document evidence and evidence fabrication. These strategies involve some potentially unethical behavior of preparers, and fall outside the framework here (based on reasoned argument and evidence framing). Rich et al. (1997) also identify evidence order as a persuasion strategy. However, because it is not clear whether any particular sequence (all pro then con reasons, or *vice versa*; interweaving pro and con reasons) is deemed to be more persuasive, we do not posit evidence order as a specific justification strategy. We content-analyzed the justifications from participating auditors in our study, and find no evidence that participants order the pro and con reasons in any systematic way.

The direction and manner with which this increased effort is exerted toward documenting persuasive justifications is likely to vary by the auditor's tacit managerial knowledge (henceforth, TMK), which refers to job knowledge related to managing self, others, and career (Wagner and Sternberg 1985). TMK has been associated with superior performance evaluations within the CPA firm (Tan and Libby 1997). The aspect of TMK most relevant here relates to knowledge about managing others, specifically, with respect to managing a superior's impressions of one's work quality. Preparers with greater TMK are likely to be more cognizant of the importance of managing a superior's impressions, and to be more aware of the different justification forms and their efficacies in persuading the reviewer that the documented conclusion and associated justifications are carefully thought out.

In particular, preparers with greater TMK are likely more aware of the greater need to use even more persuasive justifications (through using the justification forms described above) when the reviewer's preferences are dissimilar than when the preferences are similar.[7] This is because, in the former instance, if they choose to remain with their explicitly stated initial position despite the opposing view held by the reviewer, then these preparers are likely to anticipate that the reviewer would want to know why their position is more justifiable than that of the reviewer.[8] Even if the preparers were to adopt the position favored by the reviewer, high-TMK preparers are likely more cognizant of the need to invest more effort to employ the different justification forms to elaborate on *why* they changed from their initially *explicitly* stated task preferences—this is to impress upon the reviewer that they changed opinions through a reasoned process, and not to merely comply with the reviewer's task preferences.[9] On the other hand, in the case where the preparer-reviewer task preferences are similar, preparers may feel a smaller need to do so because their initial (and final) conclusion is favored by the reviewer.

However, we expect technical knowledge (henceforth, TK) to constrain the extent to which these high-TMK preparers are able to successfully employ more of these three justification forms (net persuasive evidence, breadth of issues, and evidence framing) when the preparer-reviewer preferences are dissimilar. Technical knowledge is needed to employ each of these justification forms (Libby and Luft 1993). Compared to preparers with higher TK, those with lower levels of TK may not be able to successfully use these justification forms even if they wanted to because they may not have the requisite TK to generate the pro and con reasons (for net persuasive evidence and evidence framing), and auditing- and client-related issues (breadth of issues).

---

[7] The argument here is that low TMK auditors have *less* but not *no* knowledge about these issues.

[8] Audit preparers may choose to remain with their initially expressed preference because they are convinced that it is the appropriate conclusion. They may also do so as an impression management technique in that superiors may value subordinates who are capable of defending a chosen outcome more than those who merely agree with what the reviewer prefers. In our study, we asked a Panel of 12 audit partners to assess the importance of various forms of actions and behavior on subordinates' career advancement prospects. They rated the ability of the subordinates to promote their ideas and convince superiors on the worth of their work (mean rating = 7.58, on a scale of 1 = extremely unimportant to 9 = extremely important) as more important than preparers either: (a) aligning with superior's convictions of how things need to be done even if it contradicts their own convictions (mean = 5.08, p = 0.009); (b) adjusting workpapers to suit reviewer styles and preferences (mean rating = 5.42, p = 0.001); and (c) being receptive to superior's preferences (mean = 6.08, p = 0.010, all two-tailed tests).

[9] Preparers who switch task preferences to be aligned with their reviewer's preferences may also be less likely to exert less effort to justify their documented conclusion (in that they know that their final task preferences are favored by the reviewer). However, this proclivity is likely mitigated by their incentive to manage their reputation by demonstrating that they switched preferences for good reasons (and not because they merely followed what the reviewer suggested).

In summary, we predict that preparers with dissimilar preparer-reviewer preferences are more likely to use the three justification forms (net persuasive evidence, breadth of issues, and evidence framing), and that this effect is largest for those with high TMK and high TK. This leads us to the following hypotheses.

H1a: Preparers will document more net persuasive evidence (pro less con reasons) to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

H1b: Preparers will document more breadth of issues to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

H1c: Preparers will engage in more evidence framing to justify their documented conclusions when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK.

III. RESEARCH METHOD

Participants

One hundred fifteen auditors from two Big 4 firms in Singapore participated in this study. They comprised 60 first-year audit staff and 55 third-year audit seniors.

Independent Variables

The study employs three independent variables: justification requirement (henceforth, JR) (requirement to justify their conclusion to a reviewer with similar or dissimilar task preferences), technical knowledge (TK), and tacit managerial knowledge (TMK). The first variable, JR, is a between-subjects variable manipulated at two levels. Participants are required to justify their audit conclusion (formed after knowing the reviewer's task preference) to a reviewer whose initial task preference (i.e., preferred conclusion) is either similar to or dissimilar from that of the participants (that is, their explicitly stated initial conclusion before knowing the reviewer's preference).

The second independent variable, TK, is a measured variable. The TK test contains five multiple choice questions, and an open-ended question requiring participants to briefly explain how an auditor would assess the adequacy of specific allowance for doubtful accounts. One of the multiple-choice questions pertains to factors to consider in evaluating accounting estimates, another pertains to analytical procedures related to the accounts receivable balance, while the remaining three questions pertain to accounting issues related to the sales and collections cycle of manufacturing clients. The questions are adapted from past CPA exams. With respect to the multiple-choice questions, TK is measured in terms of the number of correct responses obtained. The open-ended question is given the same weight as each of the multiple-choice questions. The response is considered correct if the participant provides a reasonable explanation for assessing the adequacy of allowance for doubtful accounts, and incorrect otherwise.

The third variable, TMK is measured using two-work related audit scenarios. These are developed by reference to Wagner and Sternberg (1985) and Tan and Libby (1997).

Eight options are presented under each scenario, and the rating for each option is on a nine-point scale (1 = extremely unimportant to 9 = extremely important). For the first scenario relating to managing the superior's impressions, participants are required to rate the importance of learning the preferences and styles of the superior, and being receptive to the superior's work styles and documentation preferences. For the second scenario relating to ways to progress up the firm's hierarchy, participants are required to rate the importance of strategies such as providing feedback to the superior on one's work progress, seeking feedback from the superior, socializing with the superior in non-work-related activities, drawing awareness to one's achievements, and promoting one's ideas and work.

Twelve partners (mean experience = 18 years) from two of the participating firms completed the TMK instrument. The partners' responses serve as the benchmark against which the preparers' TMK is measured.[10] TMK is measured as the sum of the squared deviations of a participant's ratings from the mean ratings of the Panel of partners. Subsequently, the squared deviation score of each participant is subtracted from the participant who obtained the highest score (that is, the worst score, 171.97), such that the higher the score, the better the participant's TMK.

Task

The doubtful accounts task is adapted from Hackenbrack and Nelson (1996) and Tan and Jamal (2001). It is set in the context of Innovate, a company that has been the client of the audit firm for the past five years. The company specializes in producing sound cards for personal computers. During the fourth quarter of the past year, the company adopts a more liberal credit policy in an attempt to meet its sales target. In order to manage the risk from its new credit policy, the company hires two experienced "credit agents" to monitor the collectability of the accounts. Hence, the financial controller feels that no additional allowance for doubtful accounts is needed. However, he is willing to provide a detailed footnote on the above matter. Currently, the company is in need of additional financing to meet its operational needs, and the bank requests an audited set of financial statements to discuss the terms of the loan. The case materials include the profit and loss account with details on sales, cost of sales, gross margin, bad debt expenses and other expenses for the current and prior years, and the debtors' aging analysis for the current year. Given this background, participants are to make an assessment on whether to permit a footnote disclosure or to ask for additional allowance for doubtful accounts.

Procedure

The experiment consisted of two stages, and was administered during in-house training sessions conducted by the participating firms, in the presence of a training manager. To enable us to track the participants across the two stages, and to induce accountability and motivation, we requested the participants to write down their names and contact numbers on the research instrument. In Stage 1, the participants were presented with two sets of material placed in separate sealed envelopes to ensure that the participants performed the task in the intended sequence. The first envelope contained the task materials. Participants first read the case details, assessed their agreement with the client's accounting treatment on a scale of 1 (definitely insist on additional allowance) to 11 (definitely agree with footnote disclosure), made a decision on whether to insist on additional allowance or permit a footnote disclosure, and then indicated their confidence in the decision.[11] After that, the participants proceeded to the second envelope containing some problem-solving ability questions (based on Tan and Kao [1999]) and background information questions.[12]

---

[10] To ensure that the items selected for inclusion in the tacit managerial knowledge measure have substantial agreement among the partners with regard to their level of importance, the partners' responses to the 16 items are first evaluated, and only items that have a range of 5 or less are selected. This results in the selection of ten items. These items have, on average, lower standard deviation (1.19) and skewness (−0.80), than the remaining 6 items (s.d. = 2.07, skewness = −1.39). This approach is similar to that adopted in Tan and Libby (1997).

Stage 2 of the experiment was conducted on a subsequent day of the participants' training. In Stage 2, the participants were presented with two sets of materials placed in separate sealed envelopes. The cover sheet on the first envelope informed the participants that an audit manager had agreed to participate in the exercise by acting as the audit manager in charge of the audit of Innovate. They were then informed of the need to write a memo to the manager justifying their decision on the accounting treatment to be used by the client. Finally, the instructions stated that, as in an audit review, an audit manager would review the memos. The first envelope contained the task with the reviewer's task preferences (i.e., whether the reviewer agreed or disagreed with the client's accounting treatment) indicated on a separate sheet. The instructions stated that the reviewer had briefly reviewed the case details and made a preliminary judgment. The participants first read the reviewer's task preferences, re-assessed the appropriateness of the client's accounting treatment with the knowledge of the reviewer's task preferences (on the same scale as that used in Stage 1), and indicated their decisions on the appropriateness of the client's accounting treatment together with their confidence in the decision. Subsequently, they wrote a memo to the reviewer on a separate sheet justifying their decision. The case materials (without the participants' responses to Stage 1) were included in the envelope for their reference. Finally, participants proceeded to another envelope that contained the manipulation check, TMK, and TK questions.

Dependent Variables

The dependent variables for H1a, H1b, and H1c are the specific forms of preparers' justifications. The first form is net persuasive evidence (H1a), which we define as the number of pro minus con reasons related to a documented conclusion. The second form, breadth of issues (H1b), relates to the issues covered in the preparers' justifications, and is measured by computing the sum of different auditing and accounting issues (e.g., internal control assessment, client related factors, financial statement analysis, industry-related factors, GAAP, and materiality) mentioned in the preparers' justification memos. We elicited these issues from an audit manager and two colleagues with auditing experience during a pre-test. The third form is evidence framing (H1c), which we identify by observing whether evidence is framed toward the preparers' conclusions (that is, by using semantics to highlight evidence consistent with their conclusions and downplay evidence inconsistent with their conclusions).[13] The Appendix provides examples of the various forms of justification.

---

[11] The participants also performed an additional ambiguous audit task of higher complexity related to hedging of foreign currency. Task order was counterbalanced, and order does not moderate our findings ($p > 0.340$). For this higher complexity task, participants had to assess whether it was appropriate for a holding company to hedge on behalf of its wholly owned subsidiary and, hence, reflect the exchange differences arising on the foreign currency liability in its foreign exchange reserves account. Participants found this task very complex, and wrote very few justifications. For example, only four participants engaged in evidence framing, and, in the dissimilar preparer-reviewer preference condition, only ten participants wrote any form of justification. This resulted in cell sizes that are too small for meaningful statistical tests on the effects of JR, TK, and TMK.

[12] Inferences relating to the hypotheses are unchanged with the inclusion of problem-solving ability as a control variable.

[13] Among those who used evidence framing, other than one participant, all engaged in only a single instance of evidence framing. Hence, we coded evidence framing as 1 if there was at least one instance of evidence framing, and 0 otherwise.

One of the authors and an independent coder with three years of auditing experience content-coded the justifications documented in the participants' memos. The coders were blind to the experimental treatments (that is, the reviewer's task preferences). The attributes coded are the number of pro and con reasons related to a documented conclusion, the audit and accounting issues identified (e.g., internal control assessment, client-related factors, financial statement analysis, industry-related factors, GAAP, and materiality), and the use of evidence framing. In coding the justification forms, disagreements between the two coders were resolved through discussion. The inter-rater agreement (Kappa coefficient) is high: 0.886 (pro and con statements), 0.930 (breadth of issues), and 0.862 (evidence framing, $p = 0.000$ in all instances).

IV. RESULTS

Preliminary Analyses and Manipulation Checks

As a manipulation check, we ask participants whether the judgment of the audit manager agreed or disagreed with the client. The majority correctly answered this question, and the responses of ten participants who failed the manipulation check are excluded from the analyses. Hence, the final sample consists of 105 usable responses.[14] We also ask participants to indicate the extent to which they expect the memos to be reviewed by an audit manager (1 = strongly disagree; 9 = strongly agree). The mean rating of 7.33 suggests that, on average, participants expect the memos to be subsequently reviewed by an audit manager.

The mean TK score is 3.66 (varying from 1 to 6), and the mean TMK score is 139.83 (varying from 0 to 166.36). Cronbach's alpha for the TMK questions is 0.79.[15] The two independent variables are not significantly correlated (Pearson $r = -0.115$, $p = 0.245$, two-tailed test).

To obtain a proxy for effort, we asked participants to indicate the time they started and ended the task. Our theory predicts that those who are in the dissimilar-preference condition are more likely to exert greater effort than those in the similar-preference condition. Consistent with our expectations, we find that the dissimilar-preference group spent more time writing the memo (mean = 8.74 minutes) than the similar-preference group (mean = 7.70 minutes, $p = 0.088$, one-tailed test).

We aggregate the analyses across firms as the results are not moderated by firm type ($p > 0.606$).[16]

Descriptive Statistics on Forms of Justification

Panels A and B in Table 1 present the descriptive statistics on the different justification forms that preparers use to justify their documented conclusion to a reviewer with either similar or dissimilar task preferences. The net persuasive evidence (pro less con arguments) ranges from a low of -1 to a high of 6, with an average of 1.63 (1.77) for preparers required to justify to a reviewer with similar (dissimilar) task preferences. The breadth of issues ranges from a low of 0 to a high of 6 issues, with an average of 2.17 (2.15) issues

---

[14] Inferences relating to the hypotheses are unchanged with the inclusion of the responses of participants who failed the manipulation check.

[15] Due to the multidimensional nature of the technical knowledge construct, we do not report any measure of reliability (see Bonner and Walker 1994).

[16] 48 (45.7 percent) of the participants favor the sufficiency of footnote disclosure, while 57 participants (54.3 percent) favor the need for additional allowance in their initial decisions. Decision type (footnote versus allowance) does not moderate our findings ($p > 0.229$).

TABLE 1
Descriptive Statistics on Forms of Justification
(Net Persuasive Evidence, Breadth of Issues, Evidence Framing)

	Minimum	Maximum	Mean	Standard Deviation
Panel A: Forms of Justification Used by Preparers Required to Justify Their Documented Conclusion to a Reviewer with Similar Task Preferences (n = 52)				
Net persuasive evidence	−1	6.00	1.63	1.21
Breadth of issues	0	6.00	2.17	0.98
Evidence framing	0	1.00	0.44	0.50
Panel B: Forms of Justification Used by Preparers Required to Justify Their Documented Conclusion to a Reviewer with Dissimilar Task Preferences (n = 53)				
Net persuasive evidence	−1	4.00	1.77	1.17
Breadth of issues	0	6.00	2.15	1.03
Evidence framing	0	1.00	0.40	0.49

Net persuasive evidence refers to the number of pro arguments less the number of con arguments related to a documented conclusion. Breadth of issues is the sum of accounting and auditing issues the participants consider in justifying their conclusions. Evidence framing refers to whether participants frame their evidence to highlight evidence consistent with the documented conclusion and downplay evidence inconsistent with the documented conclusion.

for preparers required to justify to a reviewer with similar (dissimilar) task preferences.[17] Finally, on average, 44 (40) percent of the participants engage in evidence framing when they are required to justify to a reviewer with similar (dissimilar) task preferences.

Test of Hypotheses

To test H1a, H1b, and H1c, separate regression analyses are run for each of the three justification forms, with JR, TMK, TK, as well as their interaction terms as independent variables. JR is coded as 0 when preparer-reviewer initial task preferences are similar, and 1 when dissimilar. Both TMK and TK are centered on their means to avoid multicollinearity with their interaction terms (Aiken and West 1991). Results of these regression analyses are shown in Table 2, Panels A to C.

Hypotheses 1a, 1b, and 1c predict that preparers will use more net persuasive evidence, breadth of issues, and evidence framing, respectively, when their initial preferred conclusions are dissimilar from (versus similar to) those of the reviewers, with the difference being higher for preparers with higher levels of TMK and TK. Table 2, Panel A, shows that the JR × TMK × TK term for net persuasive evidence is marginally significant ($p = 0.094$, one-tailed test), providing weak evidence of a three-way interaction.

To assess the nature of the interaction, we follow the procedure suggested in Aiken and West (1991), and test simple effects of JR at high (one standard deviation above mean) and low (one standard deviation below mean) levels of TMK and TK. As seen from Table 2, Panel D, when both TMK and TK are high, there is a significantly positive effect of JR on net persuasive evidence ($p = 0.014$, one-tailed test). The positively significant slope suggests that preparers document more pro arguments than con arguments when they are

---

[17] Participants generally identified issues relating to internal control risk (77.1 percent), trends/fluctuations in the key accounting numbers (56.2 percent), GAAP (23.8 percent), and materiality (22.9 percent).

TABLE 2
Regression Results of the Effects of Justification Requirement, Preparers' Technical Knowledge and Tacit Managerial Knowledge on Forms of Workpaper Justifications (Net Persuasive Evidence, Breadth of Issues, and Evidence Framing)

(n = 105)

Panel A: Net Persuasive Evidence as Dependent Variable

	Regression Coefficient	Standard Error	t	p-value
Intercept	1.614	0.162	9.971	0.000
JR	0.182	0.229	0.796	0.428
TK	−0.011	0.143	−0.079	0.937
TMK	−0.008	0.005	−1.490	0.139
JR × TK	0.292	0.186	1.569	0.120
JR × TMK	0.009	0.009	0.922	0.359
TK × TMK	0.000	0.004	−0.106	0.916
JR × TK × TMK	0.008	0.006	1.329	0.094*

Panel B: Breadth of Issues as Dependent Variable

	Regression Coefficient	Standard Error	t	p-value
Intercept	2.161	0.135	16.034	0.000
JR	−0.003	0.191	−0.014	0.989
TK	−0.265	0.119	−2.226	0.028
TMK	0.003	0.005	0.669	0.505
JR × TK	0.532	0.155	3.426	0.001
JR × TMK	0.000	0.008	−0.011	0.991
TK × TMK	−0.004	0.003	−1.220	0.225
JR × TK × TMK	0.008	0.005	1.766	0.041*

Panel C: Evidence Framing as Dependent Variable

	Regression Coefficient	Standard Error	Wald	p-value
Intercept	−0.367	0.332	1.225	0.268
JR	−0.072	0.439	0.027	0.870
TK	−0.095	0.287	0.110	0.740
TMK	0.044	0.018	5.782	0.016
JR × TK	0.127	0.356	0.126	0.722
JR × TMK	−0.043	0.023	3.420	0.064
TK × TMK	−0.012	0.013	0.912	0.340
JR × TK × TMK	0.009	0.015	0.369	0.272*

(*continued on next page*)

TABLE 2 (Continued)

Panel D: Tests of Simple Slopes of Justification Requirement (JR) at Different Levels of Preparers' Technical Knowledge (TK) and Tacit Managerial Knowledge (TMK)

	Net Persuasive Evidence as Dependent Variable				Breadth of Issues as Dependent Variable			
	Magnitude of Simple Slope	Std. Error	t	p-value	Magnitude of Simple Slope	Std. Error	t	p-value
High TK and High TMK	1.059	0.473	2.237	0.014*	0.961	0.394	2.438	0.009*
High TK and Low TMK	0.043	0.422	0.103	0.918	0.376	0.352	1.069	0.288
Low TK and High TMK	−0.212	0.465	−0.456	0.650	−0.970	0.387	−2.507	0.014
Low TK and Low TMK	−0.161	0.496	−0.325	0.746	−0.376	0.413	−0.911	0.365

* One-tailed test.

JR = requirement to justify to a reviewer with similar or dissimilar task preferences (0 = similar task preferences, 1 = dissimilar task preferences);

TK = technical knowledge (measured as the number of correct responses scored by the participants on the knowledge test; High TK is 1 standard deviation above the mean TK, and Low TK is 1 standard deviation below the mean TK); and

TMK = tacit managerial knowledge (measured as the sum-of-squared deviations of participants' ratings from the mean ratings of the Panel of partners on the TMK questions. Each participant's TMK score was then subtracted from the participant with the highest score (171.97, that is the worst score). Hence, the higher the TMK score the better is the participants' TMK; High TMK is 1 standard deviation above the mean TMK, and Low TMK is 1 standard deviation below the mean TMK).

Both TK and TMK are centered on their means to avoid multicollinearity with their interaction terms.

required to justify to a reviewer with dissimilar task preferences than when they are required to justify to a reviewer with similar task preferences. There is no statistically significant effect of JR when either TK, TMK, or both TK and TMK are low ($p > 0.650$, two-tailed test). These results are consistent with H1a.[18]

Table 2, Panel B, shows a significant three-way interaction of JR, TMK, and TK on the breadth of issues ($p = 0.041$, one-tailed test). As seen from Table 2, Panel D, the results of the tests of simple effects of JR are largely similar to that observed earlier. The results show a statistically significant positive effect of JR when both TMK and TK are high ($p = 0.009$, one-tailed test), but not when TK is high and TMK low, or when both TK and TMK are low ($p > 0.288$, two-tailed test). However, contrary to expectation, when preparers possess low TK and high TMK, there is a significantly negative effect of JR on the breadth of issues documented ($p = 0.014$, two-tailed test). Preparers with low TK and high TMK document more issues when they are required to justify to a reviewer with similar task preferences than when they are required to justify to a reviewer with dissimilar task preferences.[19] These results are partially consistent with H1b.[20]

---

[18] We separately analyze the pro and con reasons, and find a three-way interaction for pro reasons ($p = 0.042$). For con reasons, the only statistically significant effect we find is a positive main effect of TMK ($p = 0.040$).

Table 2, Panel C, presents the results of the logistic regression analysis on the presence or absence of evidence framing. Contrary to expectations, we find no statistically significant three-way interaction effect of JR × TMK × TK ($p = 0.272$, one-tailed test). Instead, we find a marginally significant two-way JR × TMK interaction ($p = 0.064$, two-tailed test). Further analysis reveals a marginally significant negative effect of JR when TMK is high ($p = 0.069$, two-tailed test) but not when TMK is low ($p = 0.219$, two-tailed test). High-TMK preparers are more likely to use evidence framing when they are required to justify to a reviewer with *similar* task preferences than when they are required to justify to a reviewer with dissimilar task preferences. TK does not moderate this effect, nor does it interact with any other variable. The absence of a TK effect may be related to the fact that among those who engaged in evidence framing, all but one engaged in only a single instance of evidence framing (i.e., involving only one pair of pro and con reasons, but not more than that). Thus, the level of TK required to generate the single pair of pro and con reason may not be high, and most participants likely have the requisite baseline knowledge.

The directional effect of JR—high-TMK preparers use more evidence framing in a similar task preference JR condition than in a dissimilar task preference JR condition—is also opposite to what we anticipated and found for the other two justification forms. We speculate that one reason is that high-TMK preparers may view evidence framing as an efficient yet effective manner to convince a similar-preference reviewer on the quality of their justifications.[21] High-TMK preparers may be aware of the need to show a two-sided argument involving both pro and con reasons (Toulmin et al. 1984; Golden and Alpert 1987), but may also be aware that with respect to a similar-preference reviewer, they may be able to economize on the extent to which they need to do so. Therefore, they conserve their effort toward efficiently using language to frame evidence that involves downplaying a single con reason through emphasizing a single pro reason (rather than engage in enumeration of a range of net persuasive evidence and breadth of issues). Consistent with this "efficiency" perspective, recall that literally all preparers using evidence framing engaged only in one (not multiple) instance of evidence framing. Also, within the similar-preference group, the use of evidence framing is negatively associated with documentation of net persuasive evidence (Pearson $r = -0.571$, $p = 0.000$, two-tailed test). This is consistent with the argument that evidence framing is viewed as a substitute for net persuasive evidence.[22] These findings suggest that preparers strategically use different justification forms in reaction to justification requirements, conditional on the skill sets they possess. Overall, our predictions are generally supported for two (net persuasive evidence and breadth of issues) out of the three forms of justifications.

---

[19] It could be that these high-TMK preparers perceive that, given their low TK, they would be less able to generate sufficient breadth of issues needed to persuade a reviewer with dissimilar task preferences (who might require substantial breadth of issues to be persuaded). Thus, they focus their effort on a reviewer with similar task preferences. It may also be that TMK compensates for the lack of TK in this situation.

[20] In coding breadth of issues documented, different aspects of the same issue (e.g., financial statement analysis) are coded as one issue, even though these would have been considered separate pro or con reasons. To assess the sensitivity of our results to this coding, we re-analyze the results using the sum of pro and con reasons as dependent variable, and obtain similar results as those using breadth of issues as dependent variable.

[21] We find no significant difference in confidence between participants required to justify to a reviewer with similar preferences and those required to justify to a reviewer with dissimilar preferences (means = 75.87 and 75.40, respectively, $p = 0.842$, two-tailed test). Furthermore, inferences relating to the hypotheses remain unchanged with the inclusion of participants' confidence in their decision as a control variable.

Effects of Maintaining versus Switching One's Initial Task Preferences

We also perform additional analyses on the 53 (out of 105) participants who were required to justify to a reviewer with dissimilar task preferences. Specifically, we investigate whether there are any differences in the nature and extent of justification forms documented by the participants who changed their decision to align with the reviewer's expressed task preference ("switch" group; n = 15, 28 percent) and those who remained with their original decision ("no-switch" group; n = 38, 72 percent). As expected, we observe no significant differences between the two groups in terms of the use of net persuasive evidence (means = 1.60 and 1.84 for the switch and no-switch groups, respectively, $p = 0.503$, two-tailed test) and breadth of issues (means = 2.20 and 2.13 for the switch and no-switch groups, respectively, $p = 0.829$, two-tailed test). However, we find that preparers who switch task preferences (and end up with final task preferences similar to that by the reviewer) have a marginally greater propensity to use evidence framing than those who do not switch (means = 0.60 and 0.32 for the switch and no-switch groups, respectively, $p = 0.062$, two-tailed test). This result is somewhat consistent with our finding in our main analyses indicating that evidence framing is more likely used when preparer-reviewer task preferences are similar. It is also possible that evidence framing is used as a means for these preparers to acknowledge "correction" of their initial task preference by highlighting how a pro reason (in favor of their revised decision) overrode initial concerns about a con reason.[23]

Effects of Justification Forms on Assessments of Memo Quality

Does greater documentation of each of the three justification forms lead to better assessments of memo quality? Presumably, the answer is "yes" because this is likely why different preparers engage in differential justification efforts. Auditor respondents in Emby and Gibbins' (1988) survey do perceive that justifiability of a decision is a characteristic of good judgment.

To address this question, we asked two evaluators to assess the quality of the justification memos written by our participants. The evaluators comprised one audit manager with 11 years of audit experience (including experience in the audit training department), and one professor who was a former audit manager with ten years of audit experience. Evaluators were required to rate the quality of each of the preparers' justification memos on a scale of 0 (low quality) to 10 (excellent quality). To ensure that the memo quality is unaffected by the legibility of the preparers' handwriting, the handwritten justification memos are transcribed before being assessed by the reviewers.

---

[22] In contrast, within the dissimilar-preference group, there is no statistically significant association between the use of net persuasive evidence and evidence framing ($r = -0.208$, $p = 0.136$, two-tailed test). Furthermore, the negative correlation between the above two justification forms (net persuasive evidence and evidence framing) is significantly higher in the similar-preference group than in the dissimilar-preference group ($p = 0.05$, Fisher r-to-Z transformation test). Across all participants, between net persuasive evidence and evidence framing, $r = -0.393$ ($p = 0.000$); between net persuasive evidence and breadth of issues, $r = 0.479$ ($p = 0.000$); and between evidence framing and breadth of issues, $r = 0.249$ ($p = 0.010$, all two-tailed).

[23] Within the switch group, there is no statistically significant association between the use of evidence framing and net persuasive evidence (Pearson $r = -0.302$, $p = 0.275$, two-tailed test). Furthermore, within the dissimilar-preference group, there is no significant difference in confidence levels between participants who change decisions to align with the reviewer and those who remain with their initial decisions (means = 76.67 and 74.89, respectively, $p = 0.655$, two-tailed test). Finally, within the switch group, we find no main or interactive effects of TK or TMK ($p > 0.450$, two-tailed test). These tests may lack power because of the small sample size ($n = 15$).

Because the task is an ambiguous audit task for which audit quality is primarily determined by the justifiability and defensibility of the documented conclusion, the evaluators were given explicit instructions to consider the quality of the justifications documented by the preparers in support of their stated conclusion, and not base their evaluations on the appropriateness of the documented conclusion without evaluating the quality of the justifications. Both evaluators had different initial task preferences (one favored footnote disclosure, and the other favored additional allowance). There is no significant difference in the mean ratings awarded by Evaluator A to the preparers' memos with conclusions that were in agreement (mean = 5.46) and those that were in disagreement with the evaluator's preference (mean = 5.10, p = 0.198, two-tailed test), suggesting no significant effect of conclusion congruence on these ratings. However, Evaluator B gave significantly higher ratings (p = 0.002, two-tailed test) to preparers' memos with conclusions that *disagreed* with that of the evaluator (mean = 5.93) and lower ratings to memos with conclusions that agreed with the evaluator (mean = 4.73); this is directionally opposite to what is expected if the evaluator had merely rewarded conclusion congruence. Thus, the evaluator's ratings appear to be influenced by the quality of the justifications rather than by conclusion appropriateness or congruence. The scores given by the two evaluators are significantly positively correlated (Pearson r = 0.429, p = 0.000, two-tailed test). Paired t-tests also show no significant differences in the average ratings accorded by the two evaluators to each memo (p = 0.663, two-tailed test). We use the average of both evaluators' ratings as dependent variable.

Table 3, Panel A, reports the results of the regression of the three justification forms (independent variables) on evaluated performance (dependent variable). As expected, all three forms of justifications are positively associated with evaluated performance (net persuasive evidence, p = 0.015; breadth of issues, p = 0.006; evidence framing, p = 0.001, all one-tailed). To assess whether the results hold after controlling for the effects of JR, TMK, and TK, we re-perform the analyses above with JR, TMK, and TK, along with their interactions, as control variables. As shown in Table 3, Panel B, all three forms of justification are again positively associated with evaluated performance (net persuasive evidence, p = 0.020; breadth of issues, p = 0.008; evidence framing, p = 0.001, all one-tailed).[24,25]

---

[24] Together, the results satisfy the following two conditions listed by Kenny et al. (1998) to establish that justification form mediates the effects of JR, TK, and TMK on evaluated performance: (a) JR, TK, and TMK affect justification form (which we find in the main analyses; Table 2, Panels A to C), and (b) in turn, the mediator (i.e., justification form) affects evaluated performance (Table 3, Panel B), even after controlling for the presence of JR, TK, and TMK. For condition (b), if the three-way interaction effect (JR × TK × TMK) is statistically insignificant, a total mediation occurs. In our case, the interaction is marginally significant (p = 0.052, two-tailed test), which suggests that justification form partially mediates the effects of JR, TK, and TMK on evaluated performance. We also re-perform the regression analysis with evaluated performance as dependent variable, and JR, TMK, TK, and their interactions as independent variables (i.e., without the three justification forms as independent variables). As in our test of H1a, we find a statistically significant three-way interaction (p = 0.010, two-tailed test), and a statistically significant simple main effect of JR only when both TMK and TK are high (p = 0.015, two-tailed test), but not when either TK, TMK, or both TK and TMK are low (p > 0.207, two-tailed test).

[25] The variance inflation factor (VIF) for each of the three justification forms is below 2.2, which is less than the cut-off point of 10 that is indicative of multicollinearity problems (Kutner et al. 2004).

TABLE 3
Evaluators' Assessment of Preparers' Task Performance (Memo Quality)
(n = 105)

	Regression Coefficient	Standard Error	t-value	p-value
Panel A: Regression Results of the Effects of Forms of Justification on Evaluators' Assessment of Preparers' Task Performance (Memo Quality)				
Intercept	3.457	0.301	11.479	0.000
Net persuasive evidence	0.318	0.143	2.220	0.015*
Breadth of issues	0.415	0.161	2.580	0.006*
Evidence framing	1.036	0.310	3.339	0.001*
Panel B: Regression Results of the Effects of Forms of Justification on Evaluators' Assessment of Preparers' Task Performance (Memo Quality) (Controlling for the Effects of Justification Requirement, Preparers' Technical Knowledge, and Tacit Managerial Knowledge)				
Intercept	3.255	0.336	9.694	0.000
JR	0.461	0.238	1.939	0.055
TK	−0.072	0.153	−0.470	0.639
TMK	0.000	0.006	0.029	0.977
JR × TK	−0.044	0.204	−0.218	0.828
JR × TMK	0.002	0.010	0.178	0.859
TK × TMK	−0.005	0.004	−1.441	0.153
JR × TK × TMK	0.012	0.006	1.969	0.052
Net persuasive evidence	0.304	0.145	2.089	0.020*
Breadth of issues	0.409	0.166	2.456	0.008*
Evidence framing	1.058	0.313	3.375	0.001*

* One-tailed test.

Preparers' task performance (memo quality) is assessed by two evaluators (a manager with a Big 4 accounting firm, and an academic with prior experience as an audit manager), on a scale of 0 (low quality) to 10 (excellent quality). The average of their ratings serves as the dependent variable.

V. CONCLUSION

We identify three ways by which preparers attempt to justify their documented conclusions. We provide the first empirical evidence that the form and extent of justifications vary by whether preparers are required to justify their explicitly stated conclusions to a reviewer with similar or dissimilar initial task preferences, and the preparers' tacit managerial knowledge and technical knowledge. We find that high tacit managerial knowledge preparers are strategic in terms of the circumstances under which they use particular forms of justifications. They have a greater tendency to use language persuasively in framing their evidence (i.e., evidence framing) when they are required to justify to reviewers with *similar* initial task preferences. In contrast, preparers with high tacit managerial knowledge *and* high technical knowledge have a tendency to articulate and enumerate audit evidence (e.g., listing of pro versus con reasons; breadth of issues considered) to provide logical support for the stated conclusion when they are required to justify their conclusions to reviewers with *dissimilar* initial task preferences. Thus, the nature and extent of justifications used vary by environmental factors (here, the preferences of the reviewer) and attributes of the preparer (here, the preferences of the preparer, and his/her tacit managerial knowledge and technical knowledge).

Readers should note some limitations of this study. First, the justification forms that we investigate are likely a subset of the repertoire of justification approaches that preparers use. For example, consultation has been documented to be a means of improving the justifiability of a decision (Kennedy et al. 1997), but this option was not available to our participants because discussion was not permitted in our experiment. Second, our participants are staff auditors and seniors, and their justification approaches may differ from those of managers, who may have more sophisticated ways of dealing with reviewers' preferences. Third, we consider a setting where the reviewer's preferences influence the work done by preparers. It may well be that the reverse is true—the preparers' work influences the reviewer's preferences, which in turn influences the work performed by preparers. Our study does not capture this possible interactive relation, which is an avenue for future research. Fourth, in evaluating the participants' memos, evaluators were given explicit instructions to consider only the quality of the justifications. This quality measure we obtain may not generalize to what occurs in practice because without such instructions, evaluators' ratings may also be influenced by the congruence of the documented conclusion with their preferred conclusion rather than by the justification's quality. Finally, our current study uses Singapore auditors as participants, and the results may not generalize to U.S. auditors. This concern is mitigated by the fact that our participating auditors are from the Big 4 accounting firms, all of which adopt consistent global practices. Prior studies using Singaporean auditor participants have also documented results that are consistent with those using U.S. auditors (Tan and Libby 1997; Tan and Kao 1999).

We conclude by highlighting some implications of our findings. One implication is that there are potential benefits that may accrue to the audit review process from having preparers form and explicitly state their task preferences before becoming aware of the reviewer's task preferences, especially when preparer-reviewer task preferences are dissimilar. Furthermore, if the reviewer wants to motivate the preparer into thinking harder about net persuasive evidence and breadth of issues to support the documented conclusion, the reviewer could, early in the audit process, play the role of "devil's advocate" (Schwenk and Valacich 1994). The reviewer could first find out or anticipate the preparers' task preferences, and then communicate an initial task preference that is opposite to that held by the preparer. Similarly, the reviewer could invoke a counter-argument mind-set in preparers by prodding them to consider potential alternatives (Kray and Galinsky 2003). However, for these mechanisms to be effective, preparers need to possess the requisite tacit managerial knowledge and technical knowledge.

Our finding that the panel of evaluators (comprising a senior audit manager and a former audit manager) assessed justification memos using these three justification forms to be of higher quality has practical importance. Extant literature argues that decision justifiability is a characteristic of good audit judgment (Emby and Gibbins 1988), but the relation between justifications and assessed work quality has hitherto not been established. To the extent that expert witnesses also assess justification memo quality in the same way, our results suggest that CPA firms should heed the manner with which justifications are written up, as well as their associated determinants. An important avenue for future research is to assess whether our findings related to how justification forms influence perceptions of memo quality apply to other groups such as expert witnesses and jurors.

APPENDIX
Examples of Forms of Justification
(Net Persuasive Evidence, Breadth of Issues, and Evidence Framing)

Panel A: Net Persuasive Evidence

Preparers' Conclusion: Require Additional Allowance for Doubtful Accounts

Pro arguments:

- Comparing the aging report of 1999 and 2000, 2000 sees an increase in the debts between 0–90 days.

Con arguments:

- Hired two experienced "credit agents" to monitor collectability.

Preparers' Conclusion: Footnote Disclosure is Sufficient

Pro arguments:

- Client has implemented additional controls to deal with the change in credit policy.

Con arguments:

- The credential of new debtors has decreased since they are smaller companies with lower credit ratings. Thus, risk of unrecoverability increases correspondingly.

Panel B: Breadth of Issues

Internal control and related risk assessment

- Company hired two experienced "credit agents" whose task is to monitor accounts receivable collections

Financial statement analysis:

- Company's aging accounts receivable for 91–120 days and greater than 120 days has decreased (2 percent).

Materiality

- Sales made to new customers were not material to the total sales for the year.

GAAP

- Based on the principle of prudence, suggest additional allowance.

Client-related factors:

- Historically a good client for the last 5 years, hence, may have comfort with management representation.

Industry-related factors

- Volatile nature of the industry.

Panel C: Framing Evidence to be Consistent with the Conclusion

Preparers' Conclusion: Require Additional Allowance for Doubtful Accounts

- The company is very confident about its new strategy and feels that no additional accrual is necessary and merely a footnote is enough. However, based on the principle of prudence and the volatile nature of the industry, suggest additional allowance.

Preparers' Conclusion: Footnote Disclosure is Sufficient

- Although management adopted a more liberal credit policy by selling to smaller companies with lower credit ratings, management hired experienced credit agents to monitor accounts receivable collections, and, as such, no additional provision is required.

REFERENCES

Aiken, L. S., and S. G. West. 1991. *Multiple Regression: Testing and Interpreting Interactions.* Newbury Park, CA: Sage Publications, The International Professional Publishers.

American Institute of Certified Public Accountants (AICPA). 2002. *Audit Documentation.* Statement of Auditing Standards (SAS) No. 96. New York, AICPA.

Areni, C. S. 2003. The effects of structural and grammatical variables on persuasion: An elaboration likelihood model perspective. *Psychology & Marketing* 20 (April): 349–375.

Beus, L. R. 2003. Collecting evidence for the audit malpractice case. *The Practical Litigator* 14: 7–18.

Bonner, S. E., and P. L. Walker. 1994. The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review* 69 (January): 157–178.

Brockner, J. 1992. The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review* 17 (January): 39–61.

Caldwell, D. F., and C. A. O'Reilly. 1982. Responses to failure: The effects of choice and responsibility on impression management. *Academy of Management Journal* 25 (March): 121–136.

Crowley, A. E., and W. D. Hoyer. 1994. An integrative framework for understanding two-sided persuasion. *Journal of Consumer Research* 20 (March): 561–574.

Emby, C., and M. Gibbins. 1988. Good judgment in public accounting: Quality and justification. *Contemporary Accounting Research* (Spring): 287–313.

Fox, F. V., and B. M. Staw. 1979. The trapped administrator: Effects of job insecurity and policy resistance upon commitment to a course of action. *Administrative Science Quarterly* 24 (September): 449–471.

Gibbins, M., and K. T. Trotman. 2002. Audit review: Managers' interpersonal expectations and conduct of review. *Contemporary Accounting Research* (Fall): 411–444.

Golden, L. L., and M. I. Alpert. 1987. Comparative analysis of the relative effectiveness of one- and two-sided communication for contrasting products. *Journal of Advertising* 16 (1): 18–26.

Hackenbrack, K., and M. Nelson. 1996. Auditor's incentives and their application of financial accounting standards. *The Accounting Review* 71 (January): 43–60.

Hastie, R. 1984. Causes and effects of causal attribution. *Journal of Personality and Social Psychology* 46 (1): 44–56.

Hirschberg, S. 1990. *Strategies of Argument.* New York, NY: Macmillan Publishing Company.

Hirst, D. E., L. Koonce, and P. J. Simko. 1995. Investor reactions to financial analysts' research reports. *Journal of Accounting Research* 33 (Autumn): 335–351.

Kamins, M. A., and H. Assael. 1987. Two-sided versus one-sided appeals: A cognitive perspective on argumentation, source derogation, and the effect of disconfirming trial on belief change. *Journal of Marketing Research* XXIV (February): 29–39.

Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31 (Autumn): 231–245.

———, D. N. Kleinmuntz, and M. E. Peecher. 1997. Determinants of justifiability of performance in ill-structured audit tasks. *Journal of Accounting Research* 35 (Supplement): 105–123.

Kenny, D. A., D. A. Kashy, and N. Bolger. 1998. Data analysis in social psychology. In *The Handbook of Social Psychology Volume 1,* edited by D. Gilbert, S. Fiske, and G. Lindzey, 233–265. Boston, MA: McGraw-Hill.

Koonce, L., U. Anderson, and G. Marchant. 1995. Justification of decisions in auditing. *Journal of Accounting Research* 33 (Autumn): 369–384.

Kray, L. J., and A. D. Galinsky. 2003. The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decision. *Organizational Behavior and Human Decision Processes* 91 (May): 69–81.

Kuperman, J. C. 2003. Using cognitive schema theory in the development of public relations strategy: Exploring the case of firms and financial analysts following acquisition announcements. *Journal of Public Relations Research* 15 (April): 117–150.

Kutner, M. H., C. J. Nachtsheim, and J. Neter. 2004. *Applied Linear Regression Models.* 4th edition. New York, NY: McGraw-Hill/Irwin.

Libby, R., and J. Luft. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation and environment. *Accounting, Organizations and Society* 18 (July): 425–450.

Magro, A. M. 2005. Knowledge, adaptivity, and performance in tax research. *The Accounting Review* 80 (April): 703–722.

Pechmann, C. 1992. Predicting when two-sided ads will be more effective than one-sided ads: The role of correlational and correspondent inferences. *Journal of Marketing Research* XXIX (November): 441–453.

Peecher, M. 1996. The influence of auditors' justification processes on their decisions: A cognitive model and experimental evidence. *Journal of Accounting Research* (Spring): 125–140.

Petty, R. E., J. T. Cacioppo, and K. Schumann. 1983. Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research* 10 (September): 135–146.

———, and D. T. Wegener. 1991. Thought systems, argument quality, and persuasion. *Advances in Social Cognition* 4: 147–161.

Pyszczynski, T. A., and J. Greenberg. 1981. Role of disconfirmed expectancies in the instigation of attributional processing. *Journal of Personality and Social Psychology* 40 (1): 31–38.

Ricchiute, D. N. 2004. Effects of an attorney's line of argument on accountants' expert witness testimony. *The Accounting Review* 79 (January): 221–245.

Rich, J. S., I. Solomon, and K. T. Trotman. 1997. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* 22 (May): 481–505.

Schwenk, C., and J. S. Valacich. 1994. Effects of devil's advocacy and dialectical inquiry on individuals versus groups. *Organizational Behavior and Human Decision Processes* 59 (August): 210–222.

Sedor, L. M. 2002. An explanation for unintentional optimism in analysts' earnings forecasts. *The Accounting Review* 77 (October): 731–753.

Staw, B. 1976. Knee-deep in the Big Muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance* 16 (June): 27–44.

Tan, C. E., C. A. Jubb, and K. A. Houghton. 1997. Auditor judgments: The effects of the partner's views on decision outcomes and cognitive effort. *Behavioral Research in Accounting* 9 (Supplement): 157–175.

Tan, H-T., and R. Libby. 1997. Tacit managerial knowledge versus technical knowledge as determinants of audit expertise in the field. *Journal of Accounting Research* 35 (Spring): 97–113.

———, and A. Kao. 1999. Accountability effects on auditors' performance: The influence of knowledge, problem-solving ability, and task complexity. *Journal of Accounting Research* 37 (Spring): 209–223.

———, and K. Jamal. 2001. Do auditors objectively evaluate their subordinates' work? *The Accounting Review* 76 (January): 99–110.

———, and J. Yip-Ow. 2001. Are reviewers' judgments influenced by memo structure and conclusions documented in audit work papers? *Contemporary Accounting Research* 18 (Winter): 663–678.

Tetlock, P. E. 1985a. Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior* 7: 297–332.

———. 1985b. Toward an intuitive politician model of attribution processes. In *The Self and Social Life*, edited by B. R. Schlener. New York, NY: McGraw-Hill.

———. 1992. The impact of accountability on judgment and choice: Toward a social contingency model. *Advances in Experimental Social Psychology* 25: 331–376.

———. 2002. Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review* 109 (July): 451–471.

Toulmin, S. 1958. *The Uses of Argument*. Cambridge, U.K.: Cambridge University Press.

———, R. D. Reike, and A. Janik. 1984. *Introduction to Reasoning*. New York, NY: Macmillan.

Turner, C. W. 2001. Accountability demands and the auditor's evidence search strategy: The influence of reviewer preferences and the nature of the response (belief versus action). *Journal of Accounting Research* 39 (December): 683–706.

Wagner, R. K., and R. J. Sternberg. 1985. Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology* (August): 436–458.

Waller, W. F., and W. L. Felix. 1984. The auditor and learning from experience: Some conjectures. *Accounting, Organizations and Society* 9 (3–4): 383–406.

Wilks, T. J. 2002. Predecisional distortion of evidence as a consequence of real-time audit review. *The Accounting Review* 77 (January): 51–71.

Winograd, B. N., J. S. Gerson, and B. L. Berlin. 2000. Audit practices of PricewaterhouseCoopers. *Auditing: A Journal of Practice & Theory* 19 (Fall): 175–182.

Wong, P. T. P., and B. Weiner. 1981. When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology* 40 (4): 650–663.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended December 31, 2005. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s)*. The number in parentheses reflects June–August 2005, under incoming Editor Dan Dhaliwal.

	Number of Manuscripts	Percent	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	84 (11)	18.06	84	18.06
31 ≤ Days ≤ 60	183 (43)	39.36	267	57.42
61 ≤ Days ≤ 90	140 (26)	30.11	407	87.53
91 ≤ Days ≤ 120	49 (13)	10.54	456	98.07
121 ≤ Days	09 (2)	1.93	465	100.00
Total	465 (95)	100.00		

The mean review time was 55.93; the median review time was 77 days.

New Submissions by calendar year (January 1 to December 31)

1992	241
1993	234
1994	231
1995	195
1996	230
1997	195
1998	196
1999	239
2000	260
2001	260*
2002	324
2003	327
2004	307
2005 (1/1–8/31)	267** (8 months)

* 2001—in addition to the 260 regular submissions, 68 MSs were submitted to the *TAR* Quality of Earnings Conference.
** Projects out to 396 new submissions for 2005.

The acceptance rate (defined as number of MSs accepted divided by the number of new submissions) over the past 5 years is between 12–16% per year.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary.*

FORMAT

1. All manuscripts should be printed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form.* See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says ..."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 17.168.B.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at TARUAZ@email.arizona.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Dan Dhaliwal, Senior Editor, *The Accounting Review*, The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108. The submission fee in U.S. funds is $125.00 for members and $200.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

MARCH 2006 PLACEMENT ADS

The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The ***Review*** *is published in January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

201 AMERICAN UNIVERSITY OF BEIRUT, Olayan School of Business (OSB) is seeking applications for a renewable three-year appointment in all accounting areas with a competitive salary and benefits. One or more positions are available at all three ranks. Successful applicants should possess an earned Doctorate by the starting date of September 15, 2006. Responsibilities include teaching, research, and services. The American University of Beirut (AUB), founded in 1866 and chartered in the State of New York, is the Middle East's premier university. The 70-acre campus is situated in Ras Beirut peninsula next to the Mediterranean Sea. The OSB serves 1,400 students and offers B.B.A., M.B.A., and E.M.B.A. programs. OSB is currently seeking AACSB. AUB is committed to Equal Opportunity and Equity regardless of gender, race, and ethnic background. Contact: George K. Najjar, Dean, OSB-AUB, PO Box 11-0236, Beirut, Lebanon. Voice: (9611) 352-700; Fax: (9611) 750-214; Email: gnajjar@aub.edu.lb.

202 KONKUK UNIVERSITY, School of Business is seeking a full-time Visiting Professor for a period of one or two semesters (starting March 2006/September 2006/March 2007). Applicants should possess a Ph.D. in a related field of business. The essential function of the position is teaching two undergraduate courses per semester. If the professor stays for one semester, he/she only needs to teach 15 weeks with a full-time salary for 6 months. The same rule applies for one year, i.e., teaching only 30 weeks with a full salary for 12 months. Salary ranges between U.S.$55,000/yr and $75,000/yr (negotiable). Extra benefits include Korean national medical insurance (half paid), free campus accommodation (utilities included) and moving expense support (up to $2,000). Contact: Thomas Joh, Dean of Business School, Konkuk University, School of Business, Hwayang-dong 1, Kwangjin-ku, Seoul, Korea 143-701; Email: thjoh@konkuk.ac.kr; Website: http://biz.konkukkr/english/main1.asp.

203 UNIVERSITY OF TENNESSEE, KNOXVILLE is accepting applications for a full-time, tenure-track position (Assistant or Associate Professor) to begin August 1, 2006. The position is for Information Management (Systems) within the Department of Accounting and Information Management. Accordingly, candidates should hold a doctoral degree with a specialization in accounting information systems, management information systems, or other comparable degree, or expect to complete all degree requirements by starting date. Please send applications to: Kenneth E. Anderson, Search Committee Chair, Department of Accounting and Information Management, CBA, University of Tennessee, Knoxville, Tennessee 37996-0560. Email: kea@utk.edu. UT Knoxville is an EEO/AA/Title VI/Title IX/Section 504/ADA/ADEA institution in the provision of its education and employment programs and services. The University welcomes and honors people of all races, genders, creeds, cultures, and sexual orientations, and values intellectual curiosity, pursuit of knowledge, and academic freedom and integrity.

204 VILLANOVA UNIVERSITY, Department of Accountancy invites applications at all levels for a tenure-track position. All research and teaching areas will be considered, but some priority may be given to teaching Financial Accounting or AIS. The candidate should be capable of high-quality research and teaching at both the undergraduate and graduate levels. A doctorate in Accounting is required. The Department of Accountancy's programs are AACSB-accredited and our graduates are vigorously recruited by professional services firms and financial institutions. Villanova University is a Roman Catholic university, sponsored by the Augustinian Order. An AA/EEO Employer, Villanova seeks a diverse faculty committed to scholarship, service, and especially teaching, who understand, respect, and can contribute to the University's mission and values. Interested candidates should write to: Wayne G. Bremser, Ph.D., Chairperson, Department of Accountancy, College of Commerce and Finance, Villanova University, Villanova, PA 19085-1678, or apply by email (attach *curriculum vitae*): wayne.bremser@villanova.edu.

205 UNIVERSITY OF MICHIGAN, Ross School of Business has a tenure-track position available in Accounting starting in the 2006 academic year. The application should include a research paper based on the candidate's dissertation, any other published or unpublished research, evidence of teaching experience (if any), and a *curriculum vita* that includes three references. Applicants should have primary research and teaching interests in financial accounting. Applications should be sent by mail only to: Professor Patricia Dechow, Stephen M. Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109-1234. The University of Michigan offers competitive salary and fringe benefits and is an Affirmative Action/Non-Discriminatory Employer.

206 UNIVERSITY OF MINNESOTA, Curtis L. Carlson School of Management invites applications for tenure-track faculty positions in accounting at all ranks. We offer a stimulating research environment with outstanding synergies across all accounting research paradigms. Salary and benefits are competitive. All applicants must possess a doctoral degree (or be close to completion), and must have a strong commitment to accounting research and teaching. Applicants for senior rank must have an established record of research publications in top-tier academic journals and must present evidence of excellence in teaching. Send applications, including a *curriculum vita*, samples of research papers, and at least one letter of recommendation from an established scholar to: Recruiting Conference Chair, Department of Accounting, Carlson School of Management, University of Minnesota, 321-19th Avenue South, Minneapolis, MN 55455. Review of applications begins November 15, 2005 and will continue until positions are filled. The University of Minnesota is committed to the policy that all persons shall have access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

207 GEORGETOWN UNIVERSITY, McDonough School of Business invites applications for tenure-track faculty positions at all levels for Fall 2006. Applicants must have (or be close to completing) a Ph.D. in Accounting or a related field, and have a strong commitment to scholarly research and excellent teaching. Applicants should submit a letter of interest, *curriculum vita*, recommendations, and samples of publications or working papers to: Professor Patricia Fairfield, Georgetown University, McDonough School of Business, Washington, D.C. 20057. Georgetown University is an Affirmative Action/Equal Opportunity Employer.

208 ARIZONA STATE UNIVERSITY–EAST COLLEGE Business Administration program at the Polytechnic Campus seeks candidates for a nine-month tenure-track position at the Assistant or Associate Professor level in the area of Accounting. Please go to http://www.eoaa.asu.edu/jobop/jobop.asp for more details. Applications must include a letter of intent, *curriculum vitae*, the names, addresses, and telephone numbers of three references, and evidence of recent scholarly activity. Inquiries should be directed to: Dr. Timothy J. Richards, Morrison School of Agribusiness and Resource Management, Arizona State University East, 7001 East Williams Field Road, Mesa, AZ 85212; Phone: (480) 727-1488; Fax: (480) 727-1961; Email: trichards@asu.edu. Electronic applications and supporting materials must be sent to ProfSearch@asu.edu. Arizona State University is an Equal Opportunity/Affirmative Action Employer. A background check is required for employment.

209 MIDWESTERN STATE UNIVERSITY Accounting Department invites applications for a tenure-track faculty position effective Fall 2006. A doctorate (or short-term ABD status) in accounting from an AACSB-accredited university or a terminal degree appropriate to the candidate's area of specialization and appropriate professional certification are required. A record of teaching excellence in Financial Accounting and a continuing record of scholarly research and publication are preferred. Appointment will be at a rank appropriate to qualifications. Compensation is competitive. Interested candidates should send a letter of application, current *curriculum vitae*, unofficial transcripts, and a list of three references to: Dr. Ralph Fritzsch, Chair, Department of Accounting, Dillard College of Business Administration, Midwestern State University, 3410 Taft, Wichita Falls, TX 76308. Website: http://business.mwsu.edu/accounting/index.asp; Email: ralph.fritzsch@mwsu.edu. EEO/ADA.

210 UNIVERSITY OF TENNESSEE, KNOXVILLE is accepting applications for a full-time, tenure-track position (Assistant or Associate Professor) to begin August 1, 2006. The position is for Financial Accounting with an international and capital markets emphasis, within the Department of Accounting and Information Management. Accordingly, candidates should hold a doctoral degree with a specialization in these areas, or expect to complete all degree requirements by his or her starting date. Please send applications to: Kenneth E. Anderson, Search Committee Chair, Department of Accounting and Information Management, CBA, University of Tennessee, Knoxville, TN 37996-0560. Email: kea@utk.edu. UT Knoxville is an EEO/AA/Title VI/Title IX/Section 504/ADA/ADEA institution in the provision of its education and employment programs and services. The University welcomes and honors people of all races, genders, creeds, cultures, and sexual orientations, and values intellectual curiosity, pursuit of knowledge, and academic freedom and integrity.

211 UNIVERSITY OF NOTRE DAME, Department of Accountancy invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to: Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity/Affirmative Action Employer.

212 SANTA CLARA UNIVERSITY, Department of Accounting is authorized to hire two tenure-track Assistant, Associate, or Full Professors for Fall 2007. Candidate's primary teaching interest should be Financial Accounting, Managerial Accounting, or Information Systems. Candidates for Assistant Professor should have completed/nearly completed their Ph.D. and be committed to a scholarly research program. Candidates at higher ranks should have a completed doctorate and a continuing record of nationally recognized scholarly research and quality teaching. Reviews of applications will begin October 1, 2006, but informal inquiries are welcome before that date. If you are interested in learning more about this position, contact: Dr. Michael Calegari, Department of Accounting, Leavey School of Business, Santa Clara University, Santa Clara, CA 95053-0380 (or email: mcalegari@scu.edu). SCU is a private Jesuit university located in Silicon Valley with competitive salaries, housing subsidies, and summer research support. SCU is an EEO/AA employer.

THE ACCOUNTING REVIEW
ISSN 0001-4826

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFIC

THE ACCOUNTING REVIEW

VOLUME 81, No. 3 MAY 2006

SARAH ELIZABETH McVAY
Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items

ELLA MAE MATSUMURA and JAE YONG SHIN
An Empirical Analysis of an Incentive Plan with Relative Performance Measures: Evidence from a Postal Service

ANWER S. AHMED, EMRE KILIC, and GERALD J. LOBO
Does Recognition versus Disclosure Matter? Evidence from Value-Relevance of Banks' Recognized and Disclosed Derivative Financial Instruments

ANDREW P. SCHMIDT
The Persistence, Forecasting, and Valuation Implications of the Tax Change Component of Earnings

BENJAMIN C. AYERS, JOHN (XUEFENG) JIANG, and P. ERIC YEUNG
Discretionary Accruals and Earnings Management: An Analysis of Pseudo Earnings Targets

PETER CAREY and ROGER SIMNETT
Audit Partner Tenure and Audit Quality

MADHAV V. RAJAN and RICHARD E. SAOUMA
Optimal Information Asymmetry

SCOTT A. RICHARDSON, RICHARD G. SLOAN, MARK T. SOLIMAN, and İREM TUNA
The Implications of Accounting Distortions and Growth for Accruals and Profitability

P-140

17/07/06

A QUARTERLY JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2005–2006 Executive Committee of the Association are as follows:

President	Judy D. Rayburn, University of Minnesota
President-Elect	Shyam Sunder, Yale University
Vice President	James E. Hunton, Bentley College
Vice President–Education	Nancy A. Bagranoff, Old Dominion University
Vice President–Finance	Susan F. Haka, Michigan State University
Vice President–International	Alfred Wagenhofer, Universitat Graz
Vice President–Professional Relations	D. Scott Showalter, KPMG
Vice President–Publications	Morton P. Pincus, University of California, Irvine
Vice President–Research	John C. Fellingham, The Ohio State University
Vice President–Sections and Regions	Susan V. Crosson, Santa Fe Community College
Past President	Jane F. Mutchler, Georgia State University

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive Accounting Education News, a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues are based on the selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.*
(This does not include Section journals.)

Select one Association-wide journal	U.S. $145.00	with Electronic Option $165.00
Select two Association-wide journals	U.S. $155.00	with Electronic Option $175.00
All three Association-wide journals	U.S. $165.00	with Electronic Option $185.00

Associate membership for students residing in the United States and Canada:

Select one Association-wide journal	U.S. $25.00	with Electronic Option $45.00
Select two Association-wide journals	U.S. $35.00	with Electronic Option $55.00
All three Association-wide journals	U.S. $45.00	with Electronic Option $65.00

Electronic journal option available for an additional U.S. $20.00 to above-mentioned full-membership and associate-membership dues.

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Artificial Intelligence/Emerging Technologies; Auditing; Financial Accounting and Reporting; Gender Issues in Accounting; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Teaching and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to TARUAZ@email.arizona.edu or sent to Professor Dan Dhaliwal, Senior Editor. *The Accounting Review,* The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108.

The submission fee of $125.00 for members of the AAA or $200.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published five times a year in January, March, May, July, and October by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rate of U.S. $300; membership subscription rate is $25 with discounts available; and emeritus subscription rate is $8.75. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review,* 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2006 American Accounting Association. All rights reserved. Printed in the United States of America.

THE ACCOUNTING REVIEW

A QUARTERLY JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Dan S. Dhaliwal
Editors
John E. Core
Steven J. Huddart
William R. Kinney, Jr.
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan

VOL. 81 MAY 2006 No. 3

SARAH ELIZABETH McVAY
Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items 501
ELLA MAE MATSUMURA and JAE YONG SHIN
An Empirical Analysis of an Incentive Plan with Relative Performance Measures: Evidence from a Postal Service 533
ANWER S. AHMED, EMRE KILIC, and GERALD J. LOBO
Does Recognition versus Disclosure Matter? Evidence from Value-Relevance of Banks' Recognized and Disclosed Derivative Financial Instruments 567
ANDREW P. SCHMIDT
The Persistence, Forecasting, and Valuation Implications of the Tax Change Component of Earnings 589
BENJAMIN C. AYERS, JOHN (XUEFENG) JIANG, and P. ERIC YEUNG
Discretionary Accruals and Earnings Management: An Analysis of Pseudo Earnings Targets 617
PETER CAREY and ROGER SIMNETT
Audit Partner Tenure and Audit Quality 653
MADHAV V. RAJAN and RICHARD E. SAOUMA
Optimal Information Asymmetry 677
SCOTT A. RICHARDSON, RICHARD G. SLOAN, MARK T. SOLIMAN, and İREM TUNA
The Implications of Accounting Distortions and Growth for Accruals and Profitability 713
Editorial Data 744
Editorial Policy and Style Information 745

CENTRAL LIBRARY

Editorial Staff of *The Accounting Review*
2005–2006

SENIOR EDITOR
Dan S. Dhaliwal, The University of Arizona and The University of Auckland

EDITORS
John E. Core
Steven J. Huddart
William R. Kinney, Jr.
Marlys Gascho Lipe
William E. McCarthy
Mark W. Nelson
Madhav Rajan

EDITORIAL ASSISTANT
Gail Emily Fey, The University of Arizona

American Accounting Association

EXECUTIVE DIRECTOR
Tracey E. Sutherland

COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Kathy Casper
Judy Cothern
Jan Kovarik (subcontracted)
Marcia Craig (subcontractor)

OTHER AAA STAFF
Debora Gardner
Peggy Turczyn
Renie Knights
Pat Stein
Deirdre Harris (subcontracted)
Roland LaTulip
James Szelminski
Mary Cole
Barbara Brady
Joe Barciniak

Editorial Advisory and Review Board

Shannon W. Anderson, Rice University
Anil Arya, The Ohio State University
Hollis (Holly) Ashbaugh-Skaife, University of Wisconsin–Madison
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Orie E. Barron, The Pennsylvania State University
Sudipta Basu, Emory University
Anne L. Beatty, The Ohio State University
Messod Daniel Beneish, Indiana University
Daniel A. Bens, The University of Arizona
Jennifer Blouin, University of Pennsylvania
Robert J. Bloomfield, Cornell University
James R. Boatsman, Arizona State University
Christine A. Botosan, University of Utah
Mark T. Bradshaw, Harvard University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
William M. Cready, The University of Texas at Dallas
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Ilia D. Dichev, University of Michigan
Sunil Dutta, University of California, Berkeley
Merle M. Erickson, The University of Chicago
Leslie G. Eldenburg, The University of Arizona
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
Jennifer J. Gaver, University of Georgia
Frank B. Gigler, University of Minnesota, Twin Cities Campus
Dan Givoly, The Pennsylvania State University
Cristi A. Gleason, The University of Iowa
Jonathan C. Glover, Carnegie Mellon University
Michelle Hanlon, University of Michigan
R. Lynn Hannan, Georgia State University
Rachel M. Hayes, University of Utah
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
S. Paul Hribar, Cornell University
Christopher D. Ittner, University of Pennsylvania
Bruce W. Johnson, The University of Iowa
Karla M. Johnstone, The University of Wisconsin–Madison
Kathryn Kadous, Emory University
Sanjay Kallapur, Purdue University and the Indian School of Business
Steven J. Kachelmeier, The University of Texas at Austin
Bin Ke, The Pennsylvania State University
Elizabeth K. Keating, Harvard University
Inder K. Khurana, University of Missouri–Columbia
Ronald R. King, Washington University in St. Louis
April Klein, New York University
Lisa L. Koonce, The University of Texas at Austin
Ranjani Krishnan, Michigan State University
Linda K. Krull, The University of Texas at Austin
Christian Leuz, University of Pennsylvania
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Robert C. Lipe, University of Oklahoma
Oliver Zhen Li, University of Notre Dame
Laureen A. Maines, Indiana University
Carol A. Marquardt, New York University
Dawn A. Matsumoto, University of Washington
Edward L. Maydew, The University of North Carolina at Chapel Hill
Richard Mendenhall, University of Notre Dame
Michael B. Mikhail, Duke University
Gregory S. Miller, Harvard University
Lillian F. Mills, The University of Arizona
Partha Mohanram, Columbia University
Donald Moser, University of Pittsburgh
James N. Myers, Texas A&M University
Venky Nagar, University of Michigan
Karen K. Nelson, Rice University

Kaye J. Newberry, University of Houston
Doron Nissim, Columbia University
Patricia C. O'Brien, University of Waterloo
Per M. Olsson, Duke University
Mark E. Peecher, University of Illinois at Urbana–Champaign
Kathy R. Petroni, Michigan State University
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, The University of Chicago
Marlene A. Plumlee, University of Utah
Gordon S. Potter, Cornell University
Grace Pownall, Emory University
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shivaram Rajgopal, University of Washington
Scott A. Richardson, University of Pennsylvania
Darren T. Roulstone, The University of Chicago
Rachel Schwartz, University of Illinois at Urbana–Champaign
K. (Shiva) Sivaramakrishnan, University of Houston
Richard G. Sloan, University of Michigan
Geoffrey B. Sprinkle, Indiana University
Phillip C. Stocken, Dartmouth College
Hun-Tong Tan, Nanyang Technology University/Singapore
Robert H. Trezevant, University of Southern California
Mark Trombley, The University of Arizona
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Richard H. Willis, Tulane University
Peter D. Wysocki, Massachusetts Institute of Technology
Hong Xie, University of Illinois at Urbana–Champaign
Xiao-Jun Zhang, University of California, Berkeley

PAST EDITORS

1926–1929	William A. Paton	1973–1975	Thomas F. Keller
1929–1943	Eric L. Kohler	1976–1978	Don T. DeCoster
1944–1947	A. C. Littleton	1978–1983	Stephen A. Zeff
1948–1949	Robert L. Dixon	1983–1987	Gary L. Sundem
1950–1959	Frank P. Smith	1987–1990	William R. Kinney, Jr.
1960–1962	Robert K. Mautz	1990–1994	A. Rashad Abdel-khalik
1963–1964	Lawrence L. Vance	1994–1997	Robert P. Magee
1965–1967	Wendell Trumbull	1997–2000	Gerald L. Salamon
1968–1970	Charles H. Griffin	2000–2002	Linda Smith Bamber
1971–1972	Eldon S. Hendriksen	2002–2005	Terry Shevlin

THE ACCOUNTING REVIEW
Vol. 81 No. 3
2006
pp. 501–531

Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items

Sarah Elizabeth McVay
New York University

ABSTRACT: This paper examines the classification of items *within* the income statement as an earnings management tool. Evidence is consistent with managers opportunistically shifting expenses from core expenses (cost of goods sold and selling, general, and administrative expenses) to special items. This vertical movement of expenses does not change bottom-line earnings, but overstates "core" earnings. In addition, it appears that managers use this earnings management tool to meet the analyst forecast earnings benchmark, as special items tend to be excluded from both pro forma and analyst earnings definitions.

Keywords: *earnings management; earnings components; special items; analyst forecasts.*

Data Availability: *Data are available from public sources identified in the paper.*

I. INTRODUCTION

Earnings management, the misrepresentation or masking of true economic performance, has been the focus of many papers. The bulk of this literature has focused on two general earnings management tools: accrual management and the manipulation of real economic activities. This paper examines a third potential earnings management tool that has been largely ignored to date: the deliberate misclassification of items *within* the income statement (herein referred to as *classification shifting*).

I argue that managers wishing to maximize reported performance might shift expenses down (or revenue up) the income statement to present a picture that is not consistent with

---

This paper is based on my dissertation at the University of Michigan. I am extremely grateful to the members of my dissertation committee, Russ Lundholm (Chair), Patty Dechow, Ilia Dichev, Lutz Kilian, and Tyler Shumway, for their time, interest, and guidance. I appreciate the constructive comments of two anonymous reviewers, Carol Anilowski, Jeffrey Doyle, Lucile Faurel, Natalie Gallery, Weili Ge, Helen Gernon, Venky Nagar, Karen Nelson, Scott Richardson, Joshua Ronen, Stephen Ryan, Cathy Shakespeare, Terry Shevlin (the editor), Doug Skinner, Mark Soliman, Irem Tuna, and the University of Michigan doctoral students. The paper also benefited from the comments of workshop participants at the 2005 Accounting and Finance Association of Australia and New Zealand Annual Meeting, the 2005 American Accounting Association Annual Meeting, Columbia University, Emory University, New York University, Northwestern University, Rice University, Stanford University, University of Chicago, University of Michigan, University of Oregon, University of Pennsylvania, University of Southern California, and University of Utah. I appreciate the research assistance of Kalin Kolev. An earlier version of this paper was entitled, "The Use of Special Items to Inflate Core Earnings."

Editor's note: This paper was accepted by Terry Shevlin.

Submitted: January 2005
Accepted: October 2005

economic reality. Classification shifting is distinct from accrual management and the manipulation of real activities in several ways. First, classification shifting does not change GAAP earnings; to the extent that financial statement users focus solely on GAAP earnings, classification shifting would be pointless. However, the individual components of the income statement are meant to be informative to financial statement users, facilitating analysis by grouping items with similar characteristics (FASB Accounting Concept No. 5). In general, the closer a line item is to sales, the more permanent this item tends to be (e.g., Lipe 1986; Fairfield et al. 1996). Furthermore, investors appear to recognize this distinction and weight individual line items within the income statement differently (e.g., Lipe 1986; Elliott and Hanna 1996; Francis et al. 1996; Davis 2002; Bradshaw and Sloan 2002).

Next, while all three methods of (income-increasing) earnings management raise expectations of future performance, accrual management and the manipulation of real activities also reduce earnings in future (or past) periods. In contrast, there is no "settling up" using classification shifting. Absent additional earnings management, next period's earnings are equal to actual earnings, rather than earnings less the cost of earnings management in the prior period, greatly reducing the cost of this earnings management tool. Finally, GAAP net income does not change, thus limiting the scrutiny of auditors and regulators (Nelson et al. 2002).

To document classification shifting, I focus on the allocation of expenses between core expenses (defined as cost of goods sold and selling, general, and administrative expenses) and special items. I posit that managers wishing to manage core earnings upward will shift expenses that should be classified as core expenses to special items.[1] As an anecdotal example of classification shifting,[2] the SEC determined that Borden, Inc., classified $192 million of marketing expenses as part of a restructuring charge when it should have been included in selling, general, and administrative expenses (Hwang 1994).[3]

For a sample of 76,901 firm-year observations from 1989 to 2003, I decompose firms' core earnings into expected and unexpected components by modeling expected core earnings, in a similar vein as the accrual model (Jones 1991). I find that unexpected core earnings (reported core earnings less predicted core earnings) is increasing in special items.[4] This association is consistent with managers classifying core expenses as special items, increasing both core earnings and income-decreasing special items.

---

[1] Alternatively, managers might shift revenue that should be classified as special items upward to be netted against core expenses. For example, IBM netted gains on asset sales against selling, general, and administrative expenses, presenting higher core earnings when instead the gains should have been broken out as special items (Bulkeley 2002).

[2] Additional examples of classification shifting include AmeriServe Food Distribution (which, prior to filing for bankruptcy, classified substantial operating expenses as restructuring charges, allegedly to mask deteriorating financial performance [Sherer 2000]); SmarTalk (which reported a one-time charge in 1997 that improperly included 1997 operating expenses, enabling SmarTalk to falsely inflate 1997 earnings before one-time charges [SEC Accounting and Auditing Enforcement Release No. 1721]; Waste Management (which netted one-time gains against current-period operating expenses [SEC Accounting and Auditing Enforcement Release No. 1405]); and Anicom [which, in 1999, charged over $7.65 million in expenses to a one-time charge, including a bank fee incurred when Anicom violated its working capital loan covenants, a note receivable with accrued interest, charges incurred for enhancement to a computerized inventory control system, and over $2.3 million in accounts receivable write-offs [http://www.sec.gov/litigation/complaints/complr17504.htm]).

[3] Classification shifting is a different use of special items to manage earnings than the interperiod shifting documented in Burgstahler et al. (2002) and Moehrle (2002), who examine accrual management *across time*, using special items and restructuring charges, respectively. These studies do not examine shifting *within* the income statement.

[4] Throughout the paper, positive special items are income-decreasing; Compustat #17 is multiplied by -1. I only examine income-decreasing special items in this paper. Classification shifting using income-increasing special items is left for future research.

While the above finding is consistent with classification shifting, it is also consistent with firms experiencing efficiency gains in the year the special item is recognized by streamlining their operations or divesting unprofitable lines of business; in doing so, these firms might experience an unexpected increase in core earnings. To distinguish between these competing explanations, I examine whether the unexpected core earnings in year *t* persist into year *t*+1. A reversal of this improvement is more consistent with temporary earnings management, while the continued presence of this improvement is more consistent with economic improvements associated with the special item. I find that the unusually high core earnings associated with special items appear to reverse in the following period.

The classification shifting documented in this paper appears to be significant economically and, thus, a viable tool for managers. Based on my model of expected core earnings, I estimate that approximately 2.2 percent of reported special items, on average, are actually current-period operating expenses that are not transitory, but are opportunistically classified as special.[5] On a per-firm-year basis across the entire sample, this translates into a mean shift of $287 thousand of recurring expenses to special items, an average of roughly one-half cent per share. For firms with income-decreasing special items of at least 5 percent of sales, the estimated mean shifted amount per firm-year is $1.66 million across 8,043 firm-years, translating into almost three cents per share. Moreover, there is ample opportunity to classification shift using special items; 31 percent of the observations in my sample recognize income-decreasing special items in any given year.

It is important to note that my inferences of classification shifting rely on a model of core earnings. This is a necessary research design choice, as special-item firms tend to be performing extremely poorly. On average, these firms have lower core earnings than the Compustat population. To document that unexpected core earnings are increasing with special items requires strong performance controls. In order to adequately control for performance, I include current-year accruals. The inclusion of this variable introduces a possible bias in the model, discussed in greater detail in Section V, and is a limitation of the model and therefore this study.

To provide additional support of my inference of classification shifting, I conduct an array of supplementary tests. First, as with any earnings management mechanism, the net benefits are expected to be greater in some settings than in others. For example, Dechow and Skinner (2000) note that earnings management will likely be greater when the action allows managers to meet the analyst forecast when they otherwise would not. Consistent with classification shifting representing an earnings management tool, I find that classification shifting is more pervasive when it allows the manager to meet the analyst forecast, especially for growth firms (Skinner and Sloan 2002). The amount shifted increases to as much as 16.2 percent of special items.

Next, I examine firms with special items in year *t*+1 and find that the "reversal" of unexpected core earnings is significantly muted for these firms, consistent with them once again having the opportunity to shift core expenses to special items and thus maintain the overstated core earnings. Third, I hand-collect the income statements and related footnotes of 190 S&P 500 firms that have income-decreasing special items of at least 5 percent of sales during 1996 to 2000. For this subsample, I split Compustat special items into two groups: those that are amenable to classification shifting (such as other unusual charges) and those that are not (such as the loss on the sale of an asset). I confirm that only those special items amenable to classification shifting are associated with the unexpectedly high core earnings in year t and the subsequent reversal of this improvement in year $t+1$. Finally, I find that classification shifting is associated with negative returns in the subsequent year, suggesting that investors are surprised when expenses that were previously excluded from core earnings recur.

---

[5] This paper does not argue that investors are naïve or easily misled. It is possible that investors are able to determine that current-period core earnings are unexpectedly high, as is documented in this paper, however the financial statement users cannot ascertain *why* the core earnings are higher than expected. Managers can simply attribute the abnormal performance to economic improvements associated with the special items. In this paper, finding evidence that the unexpectedly high core earnings in year *t* *reverses* in year *t*+1 allows for *ex post* corroboration of earnings management.

Classification shifting offers a very attractive earnings management tool. In the context of special items, classification shifting can increase pro forma earnings substantially—nearly three cents per share for firms with income-decreasing special items of at least 5 percent of sales—providing managers with a relatively low-cost tool to meet analyst forecasts.[6] This finding should be informative to analysts, investors, and regulators. Furthermore, the shifting of core expenses to special items examined in this paper is just one of many possible classification-shifting schemes. Managers likely undertake classification shifting using other accounts, such as discontinued operations or research and development—an expense that is associated with future benefits. These actions do not change bottom-line earnings, but can have a significant impact on the expectations of investors and other financial statement users.

The paper proceeds as follows. Section II provides some background and develops the hypotheses. Section III discusses the data and provides descriptive statistics. Section IV introduces the model of core earnings, and Section V describes the tests and results. Section VI concludes and offers avenues for future research.

II. MOTIVATION AND HYPOTHESES

Prior accounting research has documented two main methods of earnings management. The most commonly studied method is accrual management (e.g., Healy 1985; Jones 1991; McNichols and Wilson 1988; Rangan 1998; Teoh et al. 1998; Phillips et al. 2003). Essentially, a manager can borrow earnings from future periods, through the acceleration of revenues or deceleration of expenses, in order to improve current earnings. In addition to the cost of detection, this method of earnings management bears a one-to-one cost of earnings reduction in the future; future-period earnings will be mechanically lower by the net income that was accelerated to current earnings.[7]

A second type of earnings management can occur through the manipulation of real activities, such as providing price discounts to increase sales and cutting discretionary expenditures, such as R&D, to manage earnings (e.g., Baber et al. 1991; Dechow and Sloan 1991; Bushee 1998). Such actions can increase revenues or net income, but they are also costly. For example, cutting R&D spending to manage earnings may result in the loss of future income related to the forgone R&D opportunities. On the other hand, because the manipulation of real activities is not a GAAP violation, this earnings management tool is expected to have a lower cost of detection than accrual management.

---

[6] The evidence in this paper supports prior research on pro forma earnings that suggests managers exclude recurring expenses (Doyle et al. 2003). Moreover, while prior research suggests managers use pro forma earnings to meet the analyst forecast (e.g., Lougee and Marquardt 2004; Doyle and Soliman 2005), this paper presents evidence on *how* managers might undertake this activity.

[7] Managers can also use accrual management to overstate current-period expenses (i.e., take a "big bath"). The overstatement can be used to offset future operating expenses (Burgstahler et al. 2002), or reserves associated with restructuring charges can be reversed into income in future periods (Moehrle 2002). The cost of detection from this type of earnings management is lower than that of the more traditional accrual management. For example, Nelson et al. (2002) find that the use of reserves is a common type of earnings management that is often left unadjusted by auditors if detected. However, the benefits of this type of earnings management are also delayed because it improves *future* net income.

Again, a third potential earnings management tool is the misclassification of items *within* the income statement (classification shifting). Classification shifting bears a relatively low cost: there is no accrual that later reverses, nor are there lost revenues from forgone opportunities. Moreover, because the allocation of expenses to specific accounts can be subjective, auditors might be limited in their ability to verify the appropriate classification, and, because bottom-line income does not change, they might expend less energy on the identification or compulsory adjustments of these accounts (Nelson et al. 2002).[8]

Prior research supports the viability of classification shifting. First of all, without actually misclassifying expenses, managers have been shown to manipulate the *presentation* of the income statement in order to influence perceptions of performance. For example, Kinney and Trezevant (1997) show that managers are far more likely to break out income-decreasing special items than income-increasing special items on the face of the income statement, consistent with managers wishing to highlight the transitory nature of expenses, but not income. This strategic reporting extends to earnings press releases (Schrand and Walther 2000; Bradshaw and Sloan 2002; Bowen et al. 2005). Finally, Davis (2002) finds some evidence that managers of Internet firms gross up both revenues and costs of sales to maximize reported revenues, an especially relevant metric for these firms.

In terms of *classification*, Dye (2002) presents a model where managers attempt to secure the preferred accounting classification (within GAAP) of both real transactions (such as operating versus capital leases) and income statement classifications (such as classifying transitory gains as ordinary income). Empirically, Barnea et al. (1976) conclude that managers use their subjectivity over the classification of borderline expenses that prior to APB No. 30 could be classified as ordinary or extraordinary (such as the loss on the sale of an asset) to present smoother ordinary earnings. Givoly et al. (1999) examine the allocation of income across segments and find evidence that suggests managers shift income to the highest P/E segments, thereby maximizing expected firm value. Weiss (2001) examines the treatment of transitory earnings shocks associated with the 1993 increase in corporate tax rates, and finds that managers are more likely to highlight the income-decreasing effects (consistent with Kinney and Trezevant 1997) and offset current transitory gains with income-decreasing special items in an attempt to maximize future core earnings.[9]

The above research lends credence to the idea that the placement and presentation of expenses within the income statement is a valid earnings management tool for managers. In this paper, I posit that managers *intentionally misclassify* expenses within the income statement—a GAAP violation. Specifically, I argue that managers classify a portion of *core operating expenses* as *special items* in the year that a special item is recognized. This action can alter the perceptions of financial statement users because different income statement classifications have different information content for future earnings. Moreover, investors appear to understand the distinction and do not treat line items homogeneously.

---

[8] Each of the above earnings management mechanisms affects the actual income statement filed with the SEC. Alternatively, managers might simply exclude recurring expenses from the non-GAAP earnings numbers reported in their press releases (e.g., Doyle et al. 2003; Doyle and Soliman 2005). Clearly this mechanism has a lower cost, but it might also have a lower benefit. While including core expenses in special items offers some degree of camouflage, some pro forma exclusions can be very transparent and thereby treated as recurring by investors. For example, anecdotally, non-special-item exclusions include stock-based compensation expenses, payroll taxes on stock option exercises, and amortization costs. Along these lines, Gu and Chen (2004) find that both First Call analysts and investors appear to include some items in their core earnings calculations that managers excluded in press releases.

[9] Similarly, Cameron and Gallery (2001) find that when firms report income-increasing abnormal items, they tend to recognize offsetting income-decreasing abnormal items.

To document classification shifting, I focus on the shifting of expenses between core expenses (cost of goods sold and selling, general, and administrative expenses) and special items.[10] There are many possible misclassifications to special items. Large charges, such as those related to restructurings or mergers, offer a great deal of latitude and camouflage. For example, managers can classify normal severance charges as charges resulting from the restructuring or merger. A manager might also allocate a greater percentage of legal costs or other administrative expenses than were actually related to the restructuring or merger to the "special" fees. Classification shifting is not, however, limited to these large charges. Many unusual charges might contain misclassified core expenses. Consider Y2K expenses, which might contain the salaries of permanent information technology personnel, or a litigation gain or loss might contain day-to-day legal fees.

The appropriate categorization of expenses may not be clear-cut to auditors or other outside monitors. While the substantive procedures of auditors might identify unrecorded expenses, in this setting it is not the *recognition*, but merely the *classification*, that is in question. The documentation for expenditures might be too general for the auditor to determine the appropriate classification. Furthermore, while auditors look for abnormal fluctuations in margins and other ratios, special-item firms tend to be declining in performance (see Figure 1); while reported core earnings might be higher than actual core earnings, the reported earnings are likely still below prior-period and industry benchmarks.

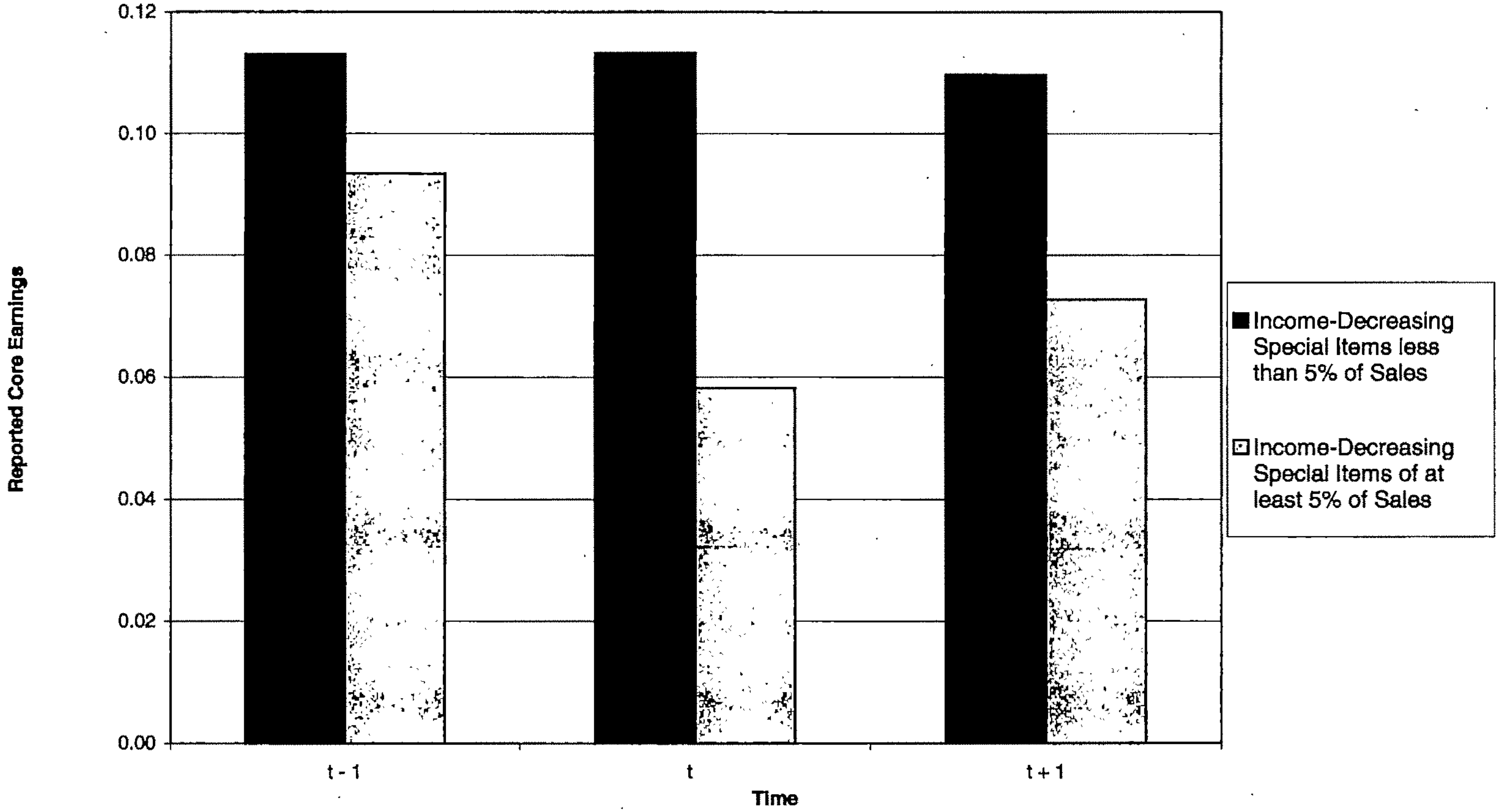
Focusing on classification shifting between core expenses and special items offers a powerful test of classification shifting for several reasons. First, core expenses and special items are clearly distinct; core expenses tend to be relatively stable, while special items are by definition unusual or infrequent. Special items have been shown to be highly transitory (e.g., Lipe 1986; Fairfield et al. 1996), and are treated accordingly by investors (e.g., Lipe 1986; Bradshaw and Sloan 2002). Therefore, classification shifting between core expenses and special items, if not fully disentangled by financial statement users, can impact expectations and thus prices. Moreover, special items tend to be excluded from core earnings by both managers (e.g., Lougee and Marquardt 2004) and analysts (e.g., Philbrick and Ricks 1991), and thus classification shifting from core expenses to special items could result in managers *ex post* meeting the analyst forecast when they otherwise would not have met this benchmark. Finally, as noted above, the shifting of expenses between core expenses and special items is viable; managers have subjectivity over the classification of expenses, and the shifting is not expected to raise red flags for outside monitors. This leads to my first hypothesis:

H1: Managers classify core expenses as special.

---

[10] Specifically, I focus on the association between core earnings and special items. Special items are material events that arise from a firm's ongoing, continuing activities, but that are either unusual in nature or infrequent in occurrence—but not both—and must be disclosed as a separate line item as part of income from continuing operations, or in footnotes to the financial statements (Revsine et al. 2005, 55). Examples of special items include (1) write-downs or write-offs of receivables, inventories, equipment, or intangibles, (2) gains or losses from the sale of equipment or investments, and (3) special one-time charges resulting from corporate restructurings. Compustat combines special items that are broken out on the income statement with those that are solely disclosed in the footnotes. However, successful classification shifting does not require special items to be highlighted on the income statement, as managers may highlight footnoted special items in the press release and/or analysts may exclude these charges from pro forma earnings. For example, in 2001, Cisco Systems Inc. reported a $2.77 billion inventory write-off, which they included in cost of goods sold within the income statement, but which was removed from pro forma earnings in their press release.

FIGURE 1
Graph of Reported Core Earnings for Large Income-Decreasing Special-Item Firms



Core Earnings is before special items and depreciation, defined as ((Sales – Cost of Goods Sold – Selling, General, and Administrative Expenses)/Sales), where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat.

In particular, I expect unexpected core earnings to be increasing in special items in year t. Furthermore, to ensure that this is due to classification shifting rather than an economic improvement associated with the special item, I expect this improvement to "reverse" in year $t+1$ as the core expenses excluded in year t recur in year $t+1$.

Clearly there are costs to classification shifting. In addition to the cost of detection, in general, managers want to avoid raising future expectations of investors or other parties. Therefore, I expect managers to classification shift to a greater degree in periods when the benefits to shifting are greater (holding costs to shifting constant). Benefits are expected to be particularly high when the earnings management allows the manager to meet earnings benchmarks (Dechow and Skinner 2000). The consensus analyst forecast has become the most important earnings benchmark in recent years (Dechow et al. 2003; Brown and Caylor 2005), and this benchmark typically excludes special items and other nonrecurring charges (Philbrick and Ricks 1991).

Moreover, meeting the analyst forecast has been shown to affect capital markets. Firms that meet analyst forecasts have an equity premium (Bartov et al. 2002; Kasznik and McNichols 2002), and missing the analyst forecast can result in a large decline in stock price, especially for high-growth firms (Skinner and Sloan 2002). Thus, the benefits to classification shifting are presumably greater for managers who can use their discretion to meet the analyst forecast, especially in high-growth firms. This leads to my second hypothesis:

> **H2:** Managers classify more core expenses as special in periods when the net benefits to classification shifting are expected to be greater.

In particular, I expect classification shifting to be more pervasive when the shifting allows managers to meet the consensus analyst forecast, especially in high-growth firms.

III. DATA, SAMPLE SELECTION, AND DESCRIPTIVE STATISTICS

Data and Sample Selection

Data are obtained for the years 1988 to 2003 from the 2003 Annual Compustat File, I/B/E/S Split-Unadjusted File, and CRSP Daily Return Tapes.[11] Each firm-year observation is required to have sufficient data to test H1. Observations with sales of less than $1 million are deleted to avoid the creation of outliers, as sales is used as a deflator for the majority of the variables. Also, firms that had a fiscal-year-end change from $t-1$ to t or from t to $t+1$ are deleted to help ensure that years are comparable. Finally, I require a minimum of 15 observations per industry per fiscal year in order to ensure a sufficiently large pool to estimate expected core earnings. Industries are classified following Fama and French (1997); results are not sensitive to the number of required observations or the industry classification scheme. The full sample has 76,901 firm-year observations.

---

[11] As my measures require one year of lagged data and one year of future data, the actual years examined are 1989–2002. The sample begins in 1988, after Compustat began reporting the newly required Cash From Operations. This allows me to calculate Accruals as Earnings less Cash From Operations as prescribed by Hribar and Collins (2002). This is particularly important in my setting because accruals are measured with error using the balance-sheet approach, especially for firms that have had Merger and Acquisition (M&A) activity (Hribar and Collins 2002), and special items often arise as a result of M&A activity. I assign a zero to special items (#17) if that data item is missing, consistent with Elliott and Hanna (1996). I also re-estimate my main analysis after excluding observations with missing special items. Results are not sensitive to this alternative approach. In addition, I assign a zero to extraordinary items and discontinued operations from the statement of cash flows (#124), which is used to calculate accruals, if that data item is missing.

Descriptive Statistics

Table 1 provides descriptive statistics for the main variables. The mean (median) core earnings for all firms, which by definition is scaled by sales, is 0.070 (0.107).[12] Mean income-decreasing special items as a percentage of sales is approximately 2.7 percent—income-increasing special items are not included in the analysis and have been set to zero.

Table 2 compares firms with and without large income-decreasing special items, where large is defined as 5 percent of sales. Not surprisingly, firms that recognize large income-decreasing special items have significantly lower core earnings than those firms that do not (−0.140 versus 0.095). Firms that undertake large write-offs or corporate restructurings tend to be poor performers (Elliott and Shaw 1988; DeAngelo et al. 1994; Carter 2000).

The change in core earnings from year $t-1$ to year t is significantly more negative for large special-item firms, suggesting that, as conditions worsen, the need for a special item such as an asset write-off or a restructuring increases. The change in core earnings from year t to $t+1$ (i.e., the year following the special item) is significantly more positive for large special-item firms (0.036 versus −0.007). This change is graphically illustrated in Figure 1 and is consistent with Brooks and Buckmaster (1976), who document that mean reversion is stronger for firms with extreme performance. Accruals are significantly more negative for large special-item firms, consistent with DeAngelo et al. (1994). Finally, the analyst forecast error is significantly more negative for large special-item firms, consistent with Abarbanell and Lehavy (2002). Table 3 provides correlations among the main variables and is discussed in the following section.

IV. MEASURING CLASSIFICATION SHIFTING

In Section II, I hypothesize that managers shift core expenses to special items; in this section, I develop a methodology to measure classification shifting. I expect core earnings of special-item firms to be overstated in the year the special item is recognized. I model the level of core earnings, and anticipate unexpected core earnings (reported core earnings less predicted core earnings) in year t to be increasing with special items in year t if managers are classification shifting. As discussed above, an alternative explanation for this association is that core earnings are unexpectedly high due to the immediate benefits of the restructuring charge or some other real economic event. In order to distinguish between real economic changes and the opportunistic behavior of managers, I examine whether the improvement associated with special items in year t reverses in year $t+1$.[13]

To test this part of the hypothesis, I model the *change* in core earnings. I expect the unexpected change in core earnings from year t to $t+1$ to be declining in special items in year t. Thus, operationally, I expect firms that classification shift to have both (1) a higher than expected level of core earnings in year t and (2) a lower than expected change in core

---

[12] Core earnings is defined as operating income before depreciation and special items scaled by sales. Depreciation is excluded to avoid the accounting effect of an asset write-down, which mechanically reduces depreciation expense. Sales is used as the scalar, rather than total assets, because assets might be systematically misstated for special-item firms. Thus, core earnings is calculated as [(Sales − Cost of Goods Sold − Selling, General, and Administrative Expenses)/Sales], where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat.

[13] As previously touched upon, if there was a real economic improvement, it is reasonable to expect the core earnings to remain at the improved level, all else equal, while mean reversion would predict a continued upward trend on the core earnings, as special-item firms tend to have core earnings that are below average (see Figure 1). On the other hand, if the unexpectedly high core earnings in the year of the special item were realized by shifting some core expenses to the special item, then special items in year t (which include core operating expenses) should be associated with a systematic decline in core earnings from year t to $t+1$, all else equal, as the previously excluded core expenses recur.

TABLE 1

Panel A: Descriptive Statistics for the Full Sample

Variable	Mean	Median	Standard Deviation	25%	75%
$SALES_t$ (in millions)	1,566.370	132.521	7,263.780	27.746	641.566
$PERCENT\ CHANGE\ IN\ SALES_{t-1,t}$	21.6%	9.0%	0.567	−1.8%	27.0%
$CORE\ EARNINGS_t$	0.070	0.107	0.385	0.039	0.196
$\Delta CORE\ EARNINGS_{t-1,t}$	0.011	0.001	0.248	−0.027	0.026
$\Delta CORE\ EARNINGS_{t,t+1}$	−0.003	0.000	0.210	−0.029	0.025
$UNEXPECTED\ CORE\ EARNINGS_t$	0.001	0.003	0.148	−0.033	0.044
$UNEXPECTED\ CHANGE\ IN\ CORE\ EARNINGS_{t,t+1}$	0.001	0.001	0.133	−0.033	0.038
$INCOME\text{-}DECREASING\ SPECIAL\ ITEMS_t$ (in millions)	13.062	0.000	79.669	0.000	0.627
$INCOME\text{-}DECREASING\ SPECIAL\ ITEMS\ AS\ A\ PERCENT\ OF\ SALES_t$	2.7%	0.0%	0.114	0.0%	0.6%
$ACCURALS_t$[a]	−0.104	−0.049	0.282	−0.124	−0.002
$ASSET\ TURNOVER\ RATIO_t$	2.82	1.93	3.36	1.06	3.22
$ANALYST\ FORECAST\ ERROR_t$	−0.016	0.000	0.093	−0.005	0.002

Panel B: Variable Definitions with Corresponding Compustat Data Item Numbers

Variable	Definition
CE_t	Core Earnings (before Special Items and Depreciation), calculated as (Sales − Cost of Goods Sold − Selling, General, and Administrative Expenses) (#13) / Sales (#12), where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat.
ΔCE_{t+1}	Change in Core Earnings, calculated as $CE_{t+1} - CE_t$.
UE_CE_t	Unexpected Core Earnings is the difference between reported and predicted Core Earnings, where the predicted value is calculated using the coefficients from model (1), as follows, estimated by fiscal year and industry and excluding firm i: $CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t$.
$UE\_\Delta CE_t$	Unexpected Change in Core Earnings is the difference between reported and predicted Change in Core Earnings, where the predicted value is calculated using the coefficients from model (2), as follows, estimated by fiscal year and industry and excluding firm i: $\Delta CE_t = \phi_0 + \phi_1 CE_{t-1} + \phi_2 \Delta CE_{t-1} + \phi_3 \Delta ATO_t + \phi_4 ACCRUALS_{t-1} + \phi_5 ACCRUALS_t + \phi_6 \Delta SALES_t + \phi_7 NEG\_\Delta SALES_t + \nu_t$.
$\%SI_t$	Income-Decreasing Special Items as a Percentage of Sales, calculated as [Special Items$_t$ (#17) × −1]/Sales$_t$(#12)[b] when Special Items are income-decreasing, and 0 otherwise.
$\Delta SALES_t$	Percent Change in Sales, calculated as (Sales$_t$ (#12) − Sales$_{t-1}$)/Sales$_{t-1}$.
$NEG\_\Delta SALES_t$	Percent Change in Sales ($\Delta SALES_t$) if $\Delta SALES_t$ is less than 0, and 0 otherwise.
$ACCRUALS_t$	Operating Accruals, calculated as [Net Income before Extraordinary Items (#123) − Cash From Operations (#308–#124)]/Sales (#12).

(*continued on next page*)

TABLE 1 (Continued)

ATO_t	Asset Turnover Ratio, defined as Sales$_t$ (#12)/((NOA$_t$ + NOA$_{t-1}$)/2), where NOA, or Net Operating Assets, is equal to the difference between Operating Assets − Operating Liabilities. Operating Assets is calculated as Total Assets (#6) less Cash (#1) and Short-Term Investments (#32). Operating liabilities is calculated as Total Assets (#6) less Total Debt (#9 and #34), less Book Value of Common and Preferred Equity (#60 and #130), less Minority Interests (#38). Average net operating assets is required to be positive.
ΔATO_t	Change in Asset Turnover, calculated as $ATO_t - ATO_{t-1}$.
FE_t	Analyst Forecast Error for year t, defined as Actual Earnings as reported by I/B/E/S less the Median I/B/E/S Analyst Forecast.
MB_t	Market-to-Book Ratio, defined as Market Value (#25 × #199) divided by Book Value (#60).

The subscript t represents the year and all variables are firm-specific. All variables are winsorized by year at the extreme 1 percent and 99 percent.

There are a maximum of 76,901 firm-year observations for each variable. Core Earnings is before special items and depreciation and defined as ((Sales − Cost of Goods Sold − Selling, General, and Administrative Expenses)/Sales), where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat. Unexpected Core Earnings in year t and Unexpected Change in Core Earnings from year t to $t+1$ are the differences between reported and predicted Core Earnings and Change in Core Earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) (shown below), estimated by fiscal year and industry and excluding firm i:

$$CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t \quad (1)$$

$$\Delta CE_t = \phi_0 + \phi_1 CE_{t-1} + \phi_2 \Delta CE_{t-1} + \phi_3 \Delta ATO_t + \phi_4 ACCRUALS_{t-1} + \phi_5 ACCRUALS_t + \phi_6 \Delta SALES_t + \phi_7 NEG\_\Delta SALES_t + \nu_t \quad (2)$$

[a] Note that accruals are scaled by sales. For comparison purposes, I also calculate accruals scaled by beginning of period assets. The mean (median) is −0.054 (−0.049), consistent with the magnitudes in prior research.

[b] Note that I multiply special items by −1 in order to be consistent with the general term special items representing an income-decreasing item.

---

earnings in year $t+1$. This prediction is opposite to what is expected to occur in the normal course of business as a result of special items. Referring to Figure 1, reported core earnings for large income-decreasing special-item firms fall, on average, in the year the special item is recognized, and improve, on average, in the next year.

I develop a model of expected core earnings, first in levels (to examine year t) and then in changes (to examine year $t+1$). This model attempts to control for economic performance as well as for macroeconomic and industry shocks. To model the level of, and change in, core earnings (CE), I estimate the following models, respectively. Regressions are estimated by industry and fiscal year:

$$\begin{aligned} CE_t = {} & \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t \\ & + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t \qquad (1) \\ \Delta CE_t = {} & \phi_0 + \phi_1 CE_{t-1} + \phi_2 \Delta CE_{t-1} + \phi_3 \Delta ATO_t + \phi_4 ACCRUALS_{t-1} \\ & + \phi_5 ACCRUALS_t + \phi_6 \Delta SALES_t + \phi_7 NEG\_\Delta SALES_t + \nu_t \qquad (2) \end{aligned}$$

where each of the variables is described below. Table 1, Panel B, presents the calculation and corresponding Compustat data item numbers of each of the variables.

TABLE 2
Descriptive Statistics for Special Item Subgroups

Variable	Firms without Income-Decreasing Special Items ≥5% of Sales (68,858 observations)		Firms with Income-Decreasing Special Items ≥5% of Sales (8,043 observations)		p-value for Statistical Difference between Firms with and without Special Items ≥5% of Sales under a:	
	Mean	Median	Mean	Median	t-test	Wilcoxon Rank Sum Test
SALES$_t$ (in millions)	1,642.390	143.642	915.509	71.279	0.001	0.001
PERCENT CHANGE IN SALES$_{t-1,t}$	21.27%	9.42%	24.1%	3.9%	0.002	0.001
CORE EARNINGS$_t$	0.095	0.111	−0.140	0.042	0.001	0.001
ΔCORE EARNINGS$_{t-1,t}$	0.013	0.002	−0.001	−0.018	0.004	0.001
ΔCORE EARNINGS$_{t,t+1}$	−0.007	0.000	0.036	0.015	0.001	0.001
UNEXPECTED CORE EARNINGS$_t$	0.000	0.003	0.005	0.012	0.100	0.001
UNEXPECTED CHANGE IN CORE EARNINGS$_{t,t+1}$	0.000	0.001	0.004	0.005	0.013	0.001
INCOME-DECREASING SPECIAL ITEMS$_t$ (in millions)	5.782	0.000	75.389	11.279	0.093	0.001
INCOME-DECREASING SPECIAL ITEMS AS A PERCENT OF SALES$_t$	0.36%	0.00%	22.2%	12.5%	0.001	0.001
ACCURALS$_t$ (scaled by Sales)	−0.070	−0.041	−0.395	−0.212	0.001	0.001
ASSET TURNOVER RATIO$_t$	2.88	1.98	2.30	1.52	0.001	0.001
ANALYST FORECAST ERROR$_t$	−0.011	0.000	−0.055	0.000	0.001	0.001

There are 76,901 firm-year observations for each variable with the exception of *ANALYST FORECAST ERROR*, which has 42,464 firm-year observations. All variables are winsorized at 1 percent and 99 percent.
See variable definitions in Table 1, Panel B.

TABLE 3
Spearman/Pearson Correlation Matrix

	$SALES_t$	$\Delta SALES_t$	CE_t	CE_{t+1}	ΔCE_t	ΔCE_{t+1}	UE_CE_t	$UE\_\Delta CE_{t+1}$	$\%SI_t$	$ACCRUALS_t$	ATO_t	FE_t
$SALES_t$	1.000	−0.038 (0.0001)	0.056 (0.0001)	0.057 (0.0001)	−0.011 (0.0017)	0.000 (0.9348)	0.013 (0.0004)	0.008 (0.0221)	−0.025 (0.0001)	0.025 (0.0001)	0.004 (0.2792)	0.035 (0.0001)
$\Delta SALES_t$	−0.037 (0.0001)	1.000	−0.111 (0.0001)	−0.111 (0.0001)	0.437 (0.0001)	0.004 (0.3167)	−0.036 (0.0001)	−0.035 (0.0001)	0.018 (0.0001)	−0.028 (0.0001)	0.092 (0.0001)	0.042 (0.0001)
CE_t	0.345 (0.0001)	0.133 (0.0001)	1.000	0.798 (0.0001)	0.035 (0.0001)	−0.306 (0.0001)	0.408 (0.0001)	−0.034 (0.0001)	−0.292 (0.0001)	0.435 (0.0001)	−0.058 (0.0001)	0.118 (0.0001)
CE_{t+1}	0.337 (0.0001)	0.071 (0.0001)	0.845 (0.0001)	1.000	−0.003 (0.3505)	0.261 (0.0001)	0.301 (0.0001)	0.310 (0.0001)	−0.227 (0.0001)	0.293 (0.0001)	−0.073 (0.0001)	0.094 (0.0001)
ΔCE_t	−0.019 (0.0001)	0.304 (0.0001)	0.178 (0.0001)	0.129 (0.0001)	1.000	−0.045 (0.0001)	0.452 (0.0001)	−0.006 (0.0887)	−0.016 (0.0001)	0.057 (0.0001)	0.040 (0.0001)	0.051 (0.0001)
ΔCE_{t+1}	−0.031 (0.0001)	−0.071 (0.0001)	−0.191 (0.0001)	0.206 (0.0001)	−0.049 (0.0001)	1.000	−0.134 (0.0001)	0.633 (0.0001)	0.107 (0.0001)	−0.224 (0.0001)	−0.024 (0.0001)	−0.050 (0.0001)
UE_CE_t	0.059 (0.0001)	−0.096 (0.0001)	0.325 (0.0001)	0.274 (0.0001)	0.451 (0.0001)	−0.075 (0.0001)	1.000	0.000 (0.9122)	0.017 (0.0001)	−0.009 (0.0099)	0.001 (0.8035)	0.049 (0.0001)
$UE\_\Delta CE_{t+1}$	0.031 (0.0001)	−0.074 (0.0001)	0.027 (0.0001)	0.254 (0.0001)	−0.005 (0.1655)	0.524 (0.0001)	0.050 (0.0001)	1.000	−0.008 (0.0250)	−0.015 (0.0001)	−0.016 (0.0001)	−0.054 (0.0001)
$\%SI_t$	0.050 (0.0001)	−0.086 (0.0001)	−0.141 (0.0001)	−0.108 (0.0001)	−0.098 (0.0001)	0.067 (0.0001)	0.014 (0.0001)	0.017 (0.0001)	1.000	−0.623 (0.0001)	−0.061 (0.0001)	−0.123 (0.0001)
$ACCRUALS_t$	0.020 (0.0001)	0.201 (0.0001)	−0.034 (0.0001)	−0.108 (0.0001)	0.093 (0.0001)	−0.154 (0.0001)	−0.150 (0.0001)	−0.047 (0.0001)	−0.268 (0.0001)	1.000	0.148 (0.0001)	0.154 (0.0001)
ATO_t	−0.025 (0.0001)	0.132 (0.0001)	−0.366 (0.0001)	−0.372 (0.0001)	0.070 (0.0001)	−0.032 (0.0001)	−0.072 (0.0001)	−0.084 (0.0001)	−0.033 (0.0001)	0.382 (0.0001)	1.000	0.024 (0.0001)
FE_t	0.093 (0.0001)	0.091 (0.0001)	0.115 (0.0001)	0.120 (0.0001)	0.183 (0.0001)	0.025 (0.0001)	0.088 (0.0001)	0.012 (0.0135)	−0.071 (0.0001)	0.049 (0.0001)	0.065 (0.0001)	1.000

There are a maximum of 76,901 firm-year observations. All variables are winsorized at 1 percent and 99 percent.
See variable definitions in Table 1, Panel B.

In the *levels* model (model 1), my first variable is lagged core earnings (CE_{t-1}). I include this variable because core earnings tends to be very persistent (note the correlation of 0.80 between core earnings and lagged core earnings in Table 3). Next, I include the asset turnover ratio (ATO_t), as it has been shown to be inversely related to profit margin (e.g., Nissim and Penman 2001), and my definition of core earnings closely parallels profit margin. Note the negative correlation between core earnings and asset turnover in Table 3, consistent with the studies referenced above. For the purpose of this paper, the inclusion of the asset turnover ratio is also important because firms that have large income-decreasing special items are likely to be making changes to their operating strategy, possibly altering their mix of margin and turnover.

Sloan (1996) finds that, holding earnings constant, accrual levels are an explanatory variable for future performance. Specifically, earnings performance attributable to the accrual component of earnings exhibits lower persistence than earnings performance attributable to the cash flow component of earnings. Thus, I include prior-year operating accruals ($ACCRUALS_{t-1}$) in my model of core earnings.

I also include current-year accruals ($ACCRUALS_t$) in my model. Extreme performance is highly correlated with changes in accrual levels (DeAngelo et al. 1994). Specifically, unusually good performance is associated with a large increase in accruals, and unusually poor performance is associated with a large decline in accruals. While it is possible that extreme accruals could be due to accrual management, this paper focuses on earnings management using special items, and therefore controlling for accruals, discretionary or otherwise, allows for a stronger prediction of core earnings.[14]

Although core earnings is scaled by sales, the relation is not expected to be constant because, as sales grow, fixed costs become smaller per sales dollar. Therefore, I include sales growth ($\Delta SALES_t$) as an explanatory variable. I also allow the slope to differ between sales increases and decreases ($NEG\_\Delta SALES_t$) because Anderson et al. (2003) find that costs increase more when activity rises than they decrease when activity falls by an equivalent amount.

To model the *change* in core earnings (model 2), I include both lagged core earnings (CE_{t-1}) and the change in core earnings from year $t-2$ to $t-1$ (ΔCE_{t-1}) to allow the model to vary the degree of mean reversion based on the prior-year's level of core earnings. This is important because mean reversion is typically more extreme in the tails (e.g., Freeman et al. 1982; Fama and French 2000). Inclusion of both levels and changes is also consistent with prior literature that forecasts changes in profitability (e.g., Fama and French 2000; Fairfield and Yohn 2001; Penman and Zhang 2002). I replace the level of asset turnover with the change in asset turnover (ΔATO_t), and retain $ACCRUALS_{t-1}$, $ACCRUALS_t$, $\Delta SALES_t$, and $NEG\_\Delta SALES_t$.

Models (1) and (2) are estimated cross-sectionally by industry and fiscal year. Table 4 provides the mean and median regression results for the model of expected core earnings. The median adjusted R^2 is quite high, at approximately 78 percent, and ranges by industry from 58 percent for Aircrafts to 94 percent for Wholesale (not tabulated).

Referring to Table 4, for the median regression, prior-year core earnings (CE_{t-1}) is a strong predictor of core earnings, as expected, with a coefficient of 0.74 and median p-value of less than 0.0001. The asset turnover ratio (ATO_t) is weakly significant (one-tailed p-value = 0.087) for the median regression, and is only significant in 36 percent of the 599 industry-year regressions. Thus, within industry-years, the asset turnover ratio does not

---

[14] While I include accruals to control performance, the inclusion of this variable results in a possible bias, which I discuss in detail in Section V.

TABLE 4
Model of Expected Core Earnings—Levels

		Dependent Variable: CE_t			
Independent Variables	**Predicted Sign**	**Mean Coefficient (one-tailed p-value)**	**Median Coefficient (one-tailed p-value)**	**Percent Significant (p-value $\leq$ 0.10, one-tailed test)**	**Percent with Sign in the Predicted Direction**
Intercept		0.04 (0.123)	0.03 (0.044)		
CE_{t-1}	+	0.74 (0.001)	0.74 (0.000)	99.7	99.8
ATO_t	−	−0.003 (0.101)	−0.002 (0.087)	36.2	66.6
$ACCRUALS_{t-1}$	−	−0.18 (0.045)	−0.18 (0.005)	70.6	84.8
$ACCRUALS_t$	+	0.22 (0.041)	0.18 (0.003)	74.0	81.5
$\Delta SALES_t$	+	0.06 (0.062)	0.05 (0.023)	58.4	78.5
$NEG\_\Delta SALES_t$	+	0.27 (0.058)	0.21 (0.019)	62.9	78.8
Adjusted R^2		75.5%	78.1%		

There are 76,901 observations and 599 industry-year regressions. Regressions are estimated by industry and fiscal year, and the p-values shown are based on one-tailed tests for each of the independent variables with the exception of the intercept, which does not have a sign prediction. p-values, rather than t-statistics, are provided due to the varying sample sizes of the specific regressions (which range from 15 to 851 observations). CE_t is Core Earnings, calculated as (Sales$_t$ − Cost of Goods Sold$_t$ − Selling, General, and Administrative Expenses$_t$)/Sales$_t$ where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat. ATO_t is the asset turnover ratio, defined as $Sales_t/((NOA_t + NOA_{t-1})/2)$, where NOA, or Net Operating Assets, is equal to Operating Assets − Operating Liabilities. Operating Assets is calculated as Total Assets − Cash and Short-Term Investments. Operating Liabilities is calculated as Total Assets − Total Debt − Book Value of Common and Preferred Equity − Minority Interests, where average net operating assets is required to be positive. $ACCRUALS_t$ is Operating Accruals, calculated as [(Net Income before Extraordinary Items − Cash From Operations)/Sales]. $\Delta SALES_t$ is the percentage change in sales from year $t-1$ to t $(Sales_t - Sales_{t-1})/(Sales_{t-1})$. $NEG\_\Delta SALES_t$ is $\Delta SALES_t$ if $\Delta SALES_t$ is negative, and 0 otherwise.
All variables are winsorized at 1 percent and 99 percent.
See Table 1, Panel B, for Compustat data item numbers.

appear to be strongly associated with core earnings. This association remains relatively weak in univariate industry-year regressions; 49 percent of the 599 regressions are statistically significant (not tabulated). These weak within-industry results are consistent with Soliman (2004), who finds that the negative relation between asset turnover and profit margin is largely driven by industry association.

Prior-year accruals ($ACCRUALS_{t-1}$) has a coefficient of −0.18, consistent with higher levels of accruals having lower earnings persistence. The positive coefficient of 0.18 on current-year accruals ($ACCRUALS_t$) is also as predicted. Consistent with Anderson et al. (2003), the slope coefficient on sales growth ($\Delta SALES_t$) is significantly larger for firms that experience a sales decline (0.05 versus 0.26, where 0.26 is obtained by summing β_5 and β_6).

Table 5 presents the mean and median regression results for Equation (2), the model of change in core earnings. The median adjusted R^2 is 51.6 percent. For the median

TABLE 5
Model of Expected Core Earnings—Changes

Independent Variables	Predicted Sign	Dependent Variable: ΔCE_t Mean Coefficient (one-tailed p-value)	Median Coefficient (one-tailed p-value)	Percent Significant (p-value ≤ 0.10, one-tailed test)	Percent with Sign in the Predicted Direction
Intercept		0.02 (0.271)	0.02 (0.121)		
CE_{t-1}	−	−0.16 (0.044)	−0.16 (0.001)	73.5	85.6
ΔCE_{t-1}	−	−0.08 (0.057)	−0.07 (0.023)	63.3	63.6
ΔATO_t	+	0.004 (0.094)	0.003 (0.074)	41.1	64.1
$ACCRUALS_{t-1}$	−	−0.17 (0.045)	−0.15 (0.007)	72.5	87.0
$ACCRUALS_t$	+	0.20 (0.039)	0.19 (0.001)	75.5	85.5
$\Delta SALES_t$	+	0.06 (0.073)	0.04 (0.044)	52.1	75.3
$NEG\_\Delta SALES_t$	+	0.25 (0.056)	0.20 (0.015)	62.6	79.6
Adjusted R^2		51.7%	51.6%		

There are 76,901 observations and 599 industry-year regressions. Regressions are estimated by industry and fiscal year, and the p-values shown are based on one-tailed tests for each of the independent variables with the exception of the intercept, which does not have a sign prediction. p-values, rather than t-statistics, are provided due to the varying sample sizes of the specific regressions (which range from 15 to 851 observations). CE_t is Core Earnings, calculated as ($Sales_t$ − Cost of Goods $Sold_t$ − Selling, General, and Administrative $Expenses_t$)/ $Sales_t$ where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat. ATO_t is the asset turnover ratio, defined as $Sales_t/((NOA_t + NOA_{t-1})/2)$, where NOA, or Net Operating Assets, is equal to Operating Assets − Operating Liabilities. Operating Assets is calculated as Total Assets − Cash and Short-Term Investments. Operating Liabilities is calculated as Total Assets − Total Debt − Book Value of Common and Preferred Equity − Minority Interests, where average net operating assets is required to be positive. $ACCRUALS_t$ is Operating Accruals, calculated as [(Net Income before Extraordinary Items − Cash From Operations)/Sales]. $\Delta SALES_t$ is the percentage change in sales from year $t-1$ to t $(Sales_t - Sales_{t-1})/(Sales_{t-1})$. $NEG\_\Delta SALES_t$ is $\Delta SALES_t$ if $\Delta SALES_t$ is negative, and 0 otherwise.
All variables are winsorized at 1 percent and 99 percent.
See Table 1, Panel B, for Compustat data item numbers.

industry-year regression, all of the variables are statistically significant (p-value < 0.10, one-tailed) and have the predicted signs. The level of core earnings is negatively associated with the change in core earnings, consistent with mean reversion (Freeman et al. 1982). The change in core earnings from year $t-2$ to $t-1$ is also negatively associated with the change in core earnings from $t-1$ to t, consistent with Brooks and Buckmaster (1976). The change in the asset turnover ratio is weakly significant in the mean and median regressions; the sign is consistent with the relation found in Penman and Zhang (2002). However, in the majority (59 percent) of the 599 industry-year regressions, this variable is not statistically significant.

Unexpected core earnings and unexpected change in core earnings are the differences between reported and predicted core earnings and change in core earnings, respectively.

The predicted values are calculated using the coefficients from models (1) and (2) above, estimated by fiscal year and industry and excluding firm *i*. Tables 1–3 provide descriptive statistics for these residuals. Referring to Table 3, core earnings and unexpected core earnings are positively correlated (0.41), as are the partitions of the discretionary accrual model. This positive association potentially confounds many studies of earnings management because both partitions (i.e., discretionary and nondiscretionary) are expected to be correlated *in the same direction* with the variable of interest (e.g., McNichols and Wilson 1988; Dechow et al. 1995; McNichols 2000). However, it is important to note that special items are *negatively* correlated with core earnings, while H1 posits a *positive* relation between special items and unexpected core earnings.

V. TEST DESIGN AND RESULTS

Main Analysis

Hypothesis 1 predicts that managers shift core expenses to special items ("classification shift"). As discussed in Section IV, if managers classification shift, then unexpected core earnings in year t is expected to be increasing with special items in year t, and the unexpected change in core earnings in year $t+1$ is expected to be decreasing with special items in year t. These anticipated associations are opposite those expected for actual core earnings, as illustrated in Figure 1. To test H1, I estimate the following regressions:

$$UE\_CE_t = \alpha_0 + \alpha_1 \%SI_t + \varepsilon_t \tag{3a}$$

$$UE\_\Delta CE_{t+1} = \eta_0 + \eta_1 \%SI_t + \nu_{t+1} \tag{3b}$$

where UE_CE_t is unexpected core earnings in year t, and $UE\_\Delta CE_{t+1}$ is unexpected change in core earnings in year $t+1$, the difference between reported and predicted core earnings and change in core earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) above, estimated by fiscal year and industry and excluding firm i. The variable $\%SI_t$ is defined as income-decreasing special items scaled by sales, both in year t. Note that a positive special item corresponds to an income-decreasing special item (income-increasing special items are set to zero); I therefore predict α_1 to be positive and η_1 to be negative. Recall that many other variables were used in the generation of unexpected core earnings, my dependent variable. As such, I do not add additional control variables to Equations (3a) and (3b). In untabulated tests I include industry-specific indicator variables; results are not sensitive to the inclusion of these variables.

I consider three samples in Table 6, (1) all Compustat firms, (2) those firms with nonzero income-decreasing special items, and (3) those firms with income-decreasing special items of at least 5 percent of sales, presented in the first, second, and third columns of results, respectively. Referring to the first column in Table 6, I find that special items are positively associated with unexpected core earnings ($\alpha_1 = 0.022$), as predicted. A one standard deviation increase in special items is expected to result in an increase in unexpected core earnings (scaled by sales) of 25 basis points [0.022 (the coefficient from Table 6) multiplied by 0.114 (the standard deviation of income-decreasing special items for this sample)]. Not surprisingly, as the sample is narrowed down to focus on those firms that have greater opportunity to classification shift, the adjusted R^2 of the models increases monotonically. For those firms with special items of at least 5 percent of sales, a one standard deviation increase in special items is expected to result in an increase in unexpected core earnings of 63 basis points.

TABLE 6
Regression of Unexpected Core Earnings and Future Unexpected Change in Core Earnings on Special Items as a Percentage of Sales

		Dependent Variable = UE_CE_t		
Independent Variables	**Predicted Sign**	**All Compustat Firms**	**Non-Zero Income-Decreasing Special Items**	**Income-Decreasing Special Items ≥ 5% of Sales**
Intercept		0.000 (0.26)	0.003 (2.18)	−0.000 (−0.07)
$\%SI_t$	+	0.022 (4.61)	0.017 (3.06)	0.022 (2.39)
Adjusted R^2		0.03%	0.04%	0.06%

		Dependent Variable = $UE\_\Delta CE_{t+1}$		
Independent Variables	**Predicted Sign**	**All Compustat Firms**	**Non-Zero Income-Decreasing Special Items**	**Income-Decreasing Special Items ≥ 5% of Sales**
Intercept		0.001 (2.45)	0.001 (5.42)	0.010 (3.47)
$\%SI_t$	−	−0.010 (−2.24)	−0.018 (−3.67)	−0.025 (−3.20)
Adjusted R^2		0.01%	0.05%	0.11%
Number of Observations		76,901	23,743	8,043
Standard Deviation of $\%SI_t$		0.114	0.193	0.286

The sample consists of 76,901 firm-year observations, t-statistics are shown in parentheses. Unexpected Core Earnings in year t (UE_CE_t) and Unexpected Change in Core Earnings from year t to $t+1$ ($UE\_\Delta CE_{t+1}$) are the differences between reported and predicted Core Earnings and Change in Core Earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) (shown below), estimated by fiscal year and industry and excluding firm i:

$$CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACCRUALS_{t-1} + \beta_4 ACCRUALS_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t \quad (1)$$

$$\Delta CE_t = \phi_0 + \phi_1 CE_{t-1} + \phi_2 \Delta CE_{t-1} + \phi_3 \Delta ATO_t + \phi_4 ACCRUALS_{t-1} + \phi_5 ACCRUALS_t + \phi_6 \Delta SALES_t + \phi_7 NEG\_\Delta SALES_t + \nu_t \quad (2)$$

CE_t is Core Earnings, calculated as (Sales$_t$ − Cost of Goods Sold$_t$ − Selling, General, and Administrative Expenses$_t$)/Sales$_t$ where Cost of Goods Sold and Selling, General, and Administrative Expenses exclude Depreciation and Amortization, as determined by Compustat. ATO_t is the asset turnover ratio, defined as $Sales_t/((NOA_t + NOA_{t-1})/2)$, where NOA is Net Operating Assets, and is defined in Table 1, Panel B. $ACCRUALS_t$ is Operating Accruals, calculated as [(Net Income before Extraordinary Items − Cash From Operations)/Sales]. $\Delta SALES_t$ is the percentage change in sales from year $t-1$ to t ($(Sales_t - Sales_{t-1})/(Sales_{t-1})$). $NEG\_\Delta SALES_t$ is $\Delta SALES_t$ if $\Delta SALES_t$ is negative, and 0 otherwise. $\%SI_t$ is income-decreasing special items as reported by Compustat (where positive specials item are income-decreasing and income-increasing special items are set to zero) scaled by sales, both in year t.

All variables are winsorized at 1 percent and 99 percent.

See Table 1, Panel B, for Compustat data item numbers.

Referring to the results for Equation (3b), presented in the lower half of Table 6, as predicted, special items in year t are negatively associated with the unexpected change in core earnings in year $t+1$ ($\eta_1 = -0.010$). This translates into a reversal of 12 basis points (-0.010×0.114) in year $t+1$ for a one standard deviation increase in $\%SI_t$ for the full sample, and of 72 basis points for those firms with special items of at least 5 percent of sales. Overall, the results are consistent with managers classifying some core expenses as special in the year a special item is recognized.

Serial Special Items

In the above analysis for the full sample, a one standard deviation increase in special items corresponds to core earnings that are 25 basis points higher than expected, but these special items only predict a "reversal" of 12 basis points in year $t+1$. It is important to note, however, that the reversal is expected to be smaller in the presence of interperiod shifting (Burgstahler et al. 2002). For example, if the manager shifted \$100 of year t operating expenses and \$25 of year $t+1$ operating expenses to the special item in year t, only \$75 of the \$100 would "recur" in year $t+1$. In addition, my tests provide a *lower bound* on the amount of expenses shifted to special items, as my model only picks up temporary shifting. Eastman-Kodak, for example, had "nonrecurring" losses in ten out of 12 years (Serwer 2002). They could simply classify the same nontransitory expenses as special *each year*. Approximately 16 percent of my sample firms recognize income-decreasing special items in both year t and year $t+1$. I expect the reversal to be lower when there is a special item in year $t+1$, because the manager can simply misclassify the expenses again. To provide evidence on how much lower the reversal is in these cases, I estimate the following regression:

$$UE\_\Delta CE_{t+1} = \omega_0 + \omega_1\%SI_t + \omega_2 SI\_IND_{t+1} + \omega_3\%SI_t \times SI\_IND_{t+1} + \varepsilon_{t+1} \qquad (4)$$

where SI_IND_{t+1} is an indicator variable that is equal to 1 if the firm has a special item in year $t+1$, and 0 otherwise. The results in Table 7 show that after controlling for serial special items (or the opportunity to classification shift in year $t+1$), the actual reversal of shifted expenses is higher when firms do not report a special item in year $t+1$ (for the full sample, $\omega_1 = -0.026$ as compared to $\eta_1 = -0.010$ in Table 6). In contrast, when the firm has a special item in year $t+1$, the reversal is zero, on average, for the full sample; $\omega_1 + \omega_3$ is not statistically different from zero.

Compustat Special Items

A limitation of the above analyses is the use of Compustat special items, which groups many types of special items together. These include items that are not susceptible to classification shifting, such as asset write-downs, and items that are more amenable to classification shifting, such as restructuring charges other than asset write-downs or merger-related costs. While I use Compustat special items in order to conduct a large sample study, I also hand-collect data for a subset of firms. If the results in Table 6 are evidence of classification shifting, then I expect the results to be stronger when only the special items amenable to misclassification are used.

To select my subsample, I identify all firms that were in the S&P 500 from 1996 to 2000 and had income-decreasing special items, as reported by Compustat, of at least 5 percent of sales, resulting in 190 firm-years. For this sample of firms, I examine 10-Ks and

TABLE 7
Regression of Future Unexpected Change in Core Earnings on Special Items as a Percentage of Sales and Repeat Special Items

		Dependent Variable = $UE\_\Delta CE_{t+1}$		
Independent Variables	**Predicted Sign**	**All Compustat Firms**	**Non-Zero Income-Decreasing Special Items**	**Income-Decreasing Special Items ≥ 5% of Sales**
Intercept		0.000	0.004	0.007
		(0.60)	(2.53)	(1.64)
$\%SI_t$	−	−0.026	−0.045	−0.053
		(−3.58)	(−5.23)	(−3.64)
SI_IND_{t+1}	?	0.003	0.004	0.008
		(2.46)	(2.12)	(1.31)
$\%SI_t \times SI\_IND_{t+1}$	+	0.022	0.039	0.036
		(2.56)	(3.66)	(2.08)
Adjusted R^2		0.02%	0.17%	0.28%
Number of Observations		76,901	23,743	8,043

The sample consists of 76,901 firm-year observations, t-statistics are shown in parentheses. Unexpected Change in Core Earnings from year t to $t+1$ ($UE\_\Delta CE_{t+1}$) is the difference between reported and predicted Change in Core Earnings, where the predicted values are calculated using the coefficients from model (2) (see Table 5), estimated by fiscal year and industry and excluding firm i. $\%SI_t$ is income-decreasing special items as reported by Compustat (where positive specials item are income-decreasing and income-increasing special items are set to zero) scaled by sales, both in year t. SI_IND_{t+1} is an indicator variable that is equal to 1 if special items in year $t+1$ are non-zero, and 0 otherwise.

All continuous variables are winsorized at 1 percent and 99 percent.

See Table 1, Panel B, for additional data definitions and Compustat data item numbers.

earnings announcements and record the type of transitory charge.[15] Next, I form two subsets of special items, those that are susceptible to classification shifting and those that are not. I consider PP&E write-offs, goodwill write-offs, and losses on asset sales to be unsusceptible to classification shifting (*%SI_NOT_SHIFTABLE*). All other special items are considered to be susceptible to classification shifting (*%SI_SHIFTABLE*). If the PP&E write-offs, goodwill write-offs, or losses on asset sales are not clearly broken out from susceptible charges, then I classify the entire charge as susceptible. I estimate the following regressions and provide the results in Table 8:

$$UE\_CE_t = \lambda_0 + \lambda_1\%SI\_SHIFTABLE_t + \lambda_2\%SI\_NOT\_SHIFTABLE_t + \varepsilon_t \quad (5a)$$

$$UE\_\Delta CE_{t+1} = \varphi_0 + \varphi_1\%SI\_SHIFTABLE_t + \varphi_2\%SI\_NOT\_SHIFTABLE_t + \upsilon_{t+1} \quad (5b)$$

Consistent with classification shifting, λ_1 is positive and significant, while φ_1 is negative and significant ($\lambda_1 = 0.27$, $\varphi_1 = -0.13$). In economic terms, a one standard deviation increase in $\%SI\_SHIFTABLE_t$ is expected to increase unexpected core earnings in year t by 243 basis points (0.27×0.09) and decrease the unexpected change in core earnings in year $t+1$ by 117 basis points (-0.13×0.09), where 0.09 is the standard deviation of

---

[15] Note that in terms of presentation, nearly every firm breaks out the charge on the income statement and in the press release, as noted in prior research (e.g., Kinney and Trezevant 1997; Bowen et al. 2005).

TABLE 8

Regression of Unexpected Core Earnings and Future Unexpected Change in Core Earnings on Special Items as a Percent of Sales, where Special Items are Classified as "Shiftable" or "Not Shiftable" Based on Type

Independent Variables	Predicted Sign	UE_CE_t	Predicted Sign	$UE\_\Delta CE_{t+1}$
Intercept		0.025 (2.69)		0.024 (2.47)
$\%SI\_SHIFTABLE_t$	+	0.268 (4.43)	−	−0.132 (−2.11)
$\%SI\_NOT\_SHIFTABLE_t$	?	0.024 (0.34)	?	0.181 (2.53)
Adjusted R^2		8.52%		5.10%
Number of Observations		190		190

The sample consists of 190 firm-year observations from 1996–2000, where the sample firms are in the S&P 500 in each corresponding year and had income-decreasing Compustat special items of at least 5 percent of sales in that year. t-statistics are shown in parentheses. Unexpected Core Earnings in year t (UE_CE_t) and Unexpected Change in Core Earnings from year t to $t+1$ ($UE\_\Delta CE_{t+1}$) are the differences between reported and predicted Core Earnings and Change in Core Earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) (see Tables 4 and 5), estimated by fiscal year and industry and excluding firm i. $\%SI\_NOT\_SHIFTABLE_t$ represents special items reported by the firm that were either an asset write-off (excluding inventory and receivable write-offs) or a loss on the sale of an asset, scaled by sales in the same year. All other special items that are reported by the firm are classified as $\%SI\_SHIFTABLE_t$ and are also scaled by sales.

All variables are winsorized at 1 percent and 99 percent.

See Table 1, Panel B, for additional data definitions and Compustat data item numbers.

$\%SI\_SHIFTABLE_t$. Also consistent with my conclusion of classification shifting, λ_2 is statistically insignificant. In other words, only special items that are susceptible to classification shifting experience unusually high core earnings in year t. Interestingly, φ_2 is positive and significant; special items that are not susceptible to classification shifting are associated with higher than expected changes in core earnings in the next year, instead of the reversal associated with special items that are susceptible to classification shifting. This result is consistent with real economic improvements as a result of special items, after controlling for the effects of classification shifting. Overall, a finer partition of special items into those that are susceptible to classification shifting and those that are not improves the power of the tests and corroborates the results from the larger Compustat sample.

Incentives to Meet the Analyst Forecast

Hypothesis 2 predicts that classification shifting will be more pervasive in periods when the firm has greater net benefits to classification shifting; specifically, when the classification shifting allows the manager to meet the analyst forecast. To test this hypothesis, I attempt to identify firms that would not have met the analyst forecast without classification shifting. I expect to find a greater degree of classification shifting by these firms than by the average firm. I classify firms that *ex post* met the analyst forecast by zero or one cent (*JUSTMET*) as those firms that would not have met the analyst forecast absent classification shifting.[16] In Table 9, I present the results for the following regressions:

$$UE\_CE_t = \gamma_0 + \gamma_1 \%SI_t + \gamma_2 JUSTMET_t + \gamma_3 \%SI \times JUSTMET_t + \varepsilon_t \quad (6a)$$

$$UE\_\Delta CE_{t+1} = \theta_0 + \theta_1 \%SI_t + \theta_2 JUSTMET_t + \theta_3 \%SI_t \times JUSTMET_t + \upsilon_{t+1} \quad (6b)$$

The results are consistent with managers classification shifting to a greater degree when doing so allows them to meet the analyst forecast. A one standard deviation increase in special items is expected to increase unexpected core earnings (scaled by sales) by an *additional* 31 basis points and have a subsequent reversal of an additional 34 basis points.[17]

TABLE 9
Regression of Unexpected Core Earnings and Future Unexpected Change in Core Earnings on Special Items as a Percent of Sales and Just Meeting the Analyst Forecast

Independent Variables	UE_CE_t	$UE\_\Delta CE_{t+1}$	UE_CE_t	$UE\_\Delta CE_{t+1}$
Intercept	−0.008 (−7.80)	0.002 (2.36)	−0.008 (−8.03)	0.002 (2.55)
$\%SI_t$	0.015 (1.95)	−0.003 (−0.41)	0.014 (1.90)	−0.003 (−0.38)
$JUSTMET_t$	0.017 (12.77)	−0.001 (−1.01)	0.017 (12.39)	−0.001 (−0.86)
$HIGH\ MB_t$			0.003 (1.93)	−0.001 (−0.89)
$\%SI_t \times JUSTMET_t$	0.027 (2.41)	−0.029 (−2.85)	−0.000 (−0.01)	−0.019 (−1.77)
$\%SI_t \times JUSTMET_t \times HIGH\ MB_t$			0.148 (6.76)	−0.053 (−2.72)
Adjusted R^2	0.48%	0.04%	0.60%	0.06%
Number of Observations	42,464	42,464	42,464	42,464

The sample consists of 42,464 firm-year observations, 20,606 firms that just met the analyst forecast, and 4,800 high market-to-book firms that just met the analyst forecast. t-statistics are shown in parentheses. Unexpected Core Earnings in year t (UE_CE_t) and Unexpected Change in Core Earnings from year t to t+1 ($UE\_\Delta CE_{t+1}$) are the differences between reported and predicted Core Earnings and Change in Core Earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) (see Tables 4 and 5), estimated by fiscal year and industry and excluding firm i. $\%SI_t$ is income-decreasing special items as reported by Compustat (where positive specials item are income-decreasing and income-increasing special items are set to zero) scaled by sales, both in year t. $JUSTMET_t$ is an indicator variable that is equal to 1 if the annual I/B/E/S analyst forecast error for year t is 0 or 1 cent, and 0 otherwise. *HIGH MB* firms are those in the highest quintile with respect to market-to-book ratio in year t.
All continuous variables are winsorized at 1 percent and 99 percent.
See Table 1, Panel B, for additional data definitions and Compustat data item numbers.

---

[16] As a second approach, in results not tabulated, I compare my estimate of classification shifting (e.g., 2.2 percent of special items) to the analyst forecast error (adjusting the forecast error to represent pre-tax dollars). If the estimate of classification shifting is greater than the forecast error, in pre-tax dollars, then I classify that firm as having used classification shifting to meet the analyst forecast. Results under this alternative identification procedure also strongly support the notion that managers use classification shifting to meet the analyst forecast, especially in high-growth firms.

[17] The incremental basis points are calculated by multiplying the coefficients on the interaction terms (0.027 and −0.029) by the standard deviation of special items for I/B/E/S firms (0.116).

Moreover, this finding is not an artifact of firm performance; if *JUSTMET* is replaced with an indicator variable representing all firms that met the analyst forecast versus those firms that "just" met the analyst forecast, then the interaction terms in both regressions are insignificant (not tabulated).

Some firms have a greater incentive than others to meet the analyst forecast. For example, Skinner and Sloan (2002) find that there is a severe price penalty for growth firms that miss the analyst forecast. I examine whether classification shifting is more pervasive for growth firms that I expect have used classification shifting to meet the analyst forecast. Following Skinner and Sloan (2002), I consider high-growth firms to be those in the highest quintile with respect to market-to-book ratio. The final two columns of results in Table 9 confirm the position that managers of high-growth firms classification shift more than the average firm manager in order to meet the analyst forecast. A one standard deviation increase in special items is expected to incrementally increase unexpected core earnings by 172 basis points and have a subsequent incremental reversal of 62 basis points for firms that are in the top quintile with respect to market-to-book ratio and that just met the analyst forecast. These findings provide strong evidence that managers classification shift to meet the analyst forecast, especially when they are managers of growth firms.

Stock Price Implications of Classification Shifting

One incentive to manage earnings is to temporarily maximize stock price (e.g., Rangan 1998; Teoh et al. 1998; Bartov and Mohanram 2004). Therefore, it is informative to determine whether investors are negatively surprised when the expenses that are shifted from core expenses in year t recur as core expenses in year $t+1$. As discussed in footnote 5, investors might be able to identify the abnormally high core earnings in year t, but cannot distinguish the origin, either a real economic improvement related to the special item or classification shifting. It is the recurrence of previously excluded expenses in year $t+1$ that *ex post* identifies the source of the unexpected core earnings. To investigate whether investors are surprised when expenses that were previously excluded from core earnings recur, I examine subsequent market-adjusted returns. Specifically, I estimate the following for I/B/E/S firms:[18]

$$RET\_1YR_{t+1} = \Psi_0 + \Psi_1 UE\_CE_t + \Psi_2 \%SI_t + \Psi_3 UE\_CE_t \times \%SI_t + Controls_t + \nu_{t+1} \tag{7}$$

where RET_1YR_{t+1} is the one-year-ahead market-adjusted, buy-and-hold return, inclusive of dividends, beginning four months after the end of fiscal year t, and continuing for one year.[19] I control for several known risk factors and anomalies, market value of equity, book-to-market ratio, beta, accruals, and momentum. I transform each of the independent variables to its scaled decile rank following Bernard and Thomas (1990). Deciles are formed monthly to avoid look-ahead bias. To control for cross-sectional correlation in the regression residual, I estimate the regression separately for each of the 14 years in my sample, and then construct t-statistics using the 14 sets of coefficient estimates, in a similar vein to Fama and MacBeth (1973).

---

[18] Because I find the amount of classification shifting increases dramatically to meet the analyst forecast, I focus on only those firms followed by I/B/E/S in this analysis to increase the power of my tests. If I consider all firms, results lose statistical significance.

[19] I do not require that firms have data for $t+1$ to be included in this test, as the worst classification shifters may simply fail during year $t+1$. In the event of a delisting, the delisting return (which is adjusted for the delisting bias documented in Shumway [1997]) is compounded with the abnormal return from the beginning of the portfolio formation to the delisting. The proceeds are then reinvested and earn the market return for the remainder of the 12-month period.

In results not tabulated, I find that Ψ_3 is negative and weakly statistically significant ($\Psi_3 = -0.096$; $t = 1.80$). This coefficient can be interpreted as the hedge return from a portfolio created to optimize on the information in the variable, after controlling for the other variables in the regression (Bernard and Thomas 1990). Thus, the hedge return on classification shifting is 9.6 percent. Moreover, these subsequent negative returns are more pervasive in the later years of the study, the time period when pro forma earnings became more of a focal point for investors (Bradshaw and Sloan 2002). One-year-ahead abnormal returns are −19.1 percent ($t = 2.01$) from 1996 to 2002. Thus, there is some evidence that investors do not fully disentangle classification shifting and are negatively surprised when previously shifted expenses recur as core expenses.[20]

Real Economic Actions as an Alternative Explanation

In this paper, evidence of classification shifting is presented by documenting that unexpected core earnings is increasing in special items in year *t*, but that this increase reverses in the following year. It is possible, however, that managers elect to cut discretionary spending, such as R&D or advertising expenditures, when their performance is extremely negative. If this cut is temporary, then the "reversal" that I document might simply be the recurrence of these expenses, which were, in fact, not made in year *t*. Thus, because special items and poor performance are correlated, it is possible that I am merely picking up these non-opportunistic cuts of discretionary expenditures rather than classification shifting.

I address this alternative explanation two ways. First, I estimate Equations (3a) and (3b) by performance quintile (not tabulated). If my findings merely reflect discretionary spending cuts in poorly performing firms, then I would expect to only find results in the lowest quintiles of performance. However, evidence of classification shifting is present throughout the performance quintiles, mitigating the likelihood that the effect is purely performance driven. Second, I control directly for changes in research and development expenditures to proxy for discretionary spending. Results are not sensitive to the inclusion of this control variable. Thus, it does not appear that discretionary spending cuts are driving the relation between core earnings and special items.

Model Specification as an Alternative Explanation

The model of core earnings developed in Section IV includes current-year accruals as an indicator of extreme performance. This variable is important because special-item firms tend to experience extreme negative performance. However, part of the accruals in year *t* may be due to the accrual portion of the year *t* special item. In the extreme, if accruals are entirely special items, then the residual from my model would be orthogonal to special items and there would be no bias. In actuality, however, accruals are comprised of both "normal" accruals and "special item" accruals. Both of these accruals are expected to be associated with performance. However, the association with performance may not be equal for both types of accruals, while my model is restricted to treat all types of accruals the same. Thus, if special item accruals are more or less associated with performance than

---

[20] Note that this finding does not conflict with prior research. Burgstahler et al. (2002) find that special items are weakly associated with *positive* future returns right around subsequent earnings announcements. Doyle et al. (2003) find that special items are not associated with one-year-ahead abnormal returns. I find that the portion of special items I suspect are, in fact, core expenses are negatively associated with future returns, after controlling for the main effect of special items, which is weakly positive ($\Psi_2 = 0.063$; $t = 1.42$).

"normal" accruals, then the residuals from my models might vary systematically with accrual special items.

In this section, I re-estimate core earnings and next year's change in core earnings under three alternate specifications, presented in the first three columns of results in Table 10. The first alternate specification (*no accruals*) excludes accruals from the models entirely. The second specification (*accruals = CE − CFO*) includes accruals before special items,

TABLE 10

Regression of Alternate Core Earnings and Future Change in Core Earnings Metrics on Special Items as a Percentage of Sales

Independent Variables	Pred. Sign	UE_CE_t No Accruals in the Model	UE_CE_t Accruals = $CE - CFO$	UE_CE_t Accruals = $NI - CFO - SI\_ACC$	UE_CE_t Accruals = $NI - CFO$
Intercept		0.006	0.004	0.002	0.000
		(11.04)	(7.54)	(2.83)	(0.26)
% SI_t	+	−0.174	−0.069	−0.032	0.022
		(−36.43)	(−16.88)	(−6.94)	(4.61)
Adjusted R^2		1.70%	0.37%	0.06%	0.03%

Independent Variables	Pred. Sign	$UE\_\Delta CE_{t+1}$ No Accruals in the Model	$UE\_\Delta CE_{t+1}$ Accruals = $CE - CFO$	$UE\_\Delta CE_{t+1}$ Accruals = $NI - CFO - SI\_ACC$	$UE\_\Delta CE_{t+1}$ Accruals = $NI - CFO$
Intercept		−0.000	−0.000	0.001	0.001
		(−0.34)	(−0.64)	(0.94)	(2.45)
% SI_t	−	0.060	0.052	0.010	−0.010
		(13.49)	(12.96)	(2.45)	(−2.24)
Adjusted R^2		0.23%	0.22%	0.01%	0.01%

The sample consists of 76,901 firm-year observations, t-statistics are shown in parentheses. Unexpected Core Earnings in year *t* (UE_CE_t) and Unexpected Change in Core Earnings from year *t* to *t*+1 ($UE\_\Delta CE_{t+1}$) are the differences between reported and predicted Core Earnings and Change in Core Earnings, respectively, where the predicted values are calculated using the coefficients from models (1) and (2) (shown below), estimated by fiscal year and industry and excluding firm *i*:

$$CE_t = \beta_0 + \beta_1 CE_{t-1} + \beta_2 ATO_t + \beta_3 ACC\_VAR_{t-1} + \beta_4 ACC\_VAR_t + \beta_5 \Delta SALES_t + \beta_6 NEG\_\Delta SALES_t + \varepsilon_t \qquad (1)$$

$$\Delta CE_t = \phi_0 + \phi_1 CE_{t-1} + \phi_2 \Delta CE_{t-1} + \phi_3 \Delta ATO_t + \phi_4 ACC\_VAR_{t-1} + \phi_5 ACC\_VAR_t + \phi_6 \Delta SALES_t + \phi_7 NEG\_\Delta SALES_t + \nu_t \qquad (2)$$

where *ACC_VAR* is equal to the following:

No Accruals in the Model:	*ACC_VAR* is set equal to zero,
Accruals = *CE* − *CFO*:	*ACC_VAR* is calculated as [(Core Earnings − Cash From Operations)/Sales],
Accruals = *NI* − *CFO* + *SI_ACC*:	*ACC_VAR* is calculated as [(Net Income before Extraordinary Items − Cash From Operations + Accrual Special Items[a])/Sales],
Accruals = *NI* − *CFO*:	*ACC_VAR* is calculated as [(Net Income before Extraordinary Items − Cash From Operations)/Sales].

[a] I roughly estimate accrual special items to be the sum of Compustat data items #213 (Loss [Gain] on sale of assets) and #217 (Funds from Operations, Other [which includes Reorganization Costs and Special Items]). This value is winsorized at net special items (Compustat data item #17) and set to 0 if net special items are income-increasing.

All variables are winsorized at 1 percent and 99 percent.

See Table 1, Panel B, for additional data definitions and Compustat data item numbers.

where accruals are equal to the difference between core earnings (*CE*) and cash from operations (*CFO*). Essentially, this treats all special items as accruals. The third specification (*accruals* = *NI* − *CFO* − *SI_ACC*) attempts to back out only those special items that are expected to be accruals, rather than all special items; accruals are equal to net income before extraordinary items (*NI*) less cash from operations (*CFO*), and accrual special items (*SI_ACC*).[21] Finally, for comparison purposes, the fourth column of results (*accruals* = *NI* − *CFO*) is the complete model.

I regress these three sets of alternate dependent variables on special items. Results for the full sample are no longer consistent with classification shifting. Rather, the performance effect appears to swamp any evidence of shifting; in each alternate setting, special items and unexpected core earnings are negatively associated, and special items and unexpected change in next year's core earnings are positively associated (consistent with mean reversion). To pursue the explanation that the model is weak as a result of inadequately controlling for performance, I re-estimate the main regression and include an indicator variable for non-negative core earnings (as these firms should be less subject to the strong performance effect). As supporting evidence, in each of the three alternative specifications I find a positive and significant coefficient in the initial amount of shifting and a negative and significant coefficient on the subsequent reversal (not tabulated). For example, using the model that excludes accruals entirely, $\alpha_1 = -0.152 + \alpha_3 = 0.193$, for a net shifting of 0.041 for firms with non-negative core earnings.

Referring to Table 11, results continue to support classification shifting when I focus on those special items that are amendable to shifting. In the first column, which does not have accruals in the model, special items that are amenable to classification shifting ("shiftable") are weakly positively associated with contemporaneous earnings ($\alpha_1 = 0.092$; $t = 1.44$; one-tailed p-value of 0.076) and there is a strong negative "reversal" in the subsequent year ($\gamma_1 = -0.263$; $t = -3.51$). In results not tabulated, if I examine only those firms with non-negative core earnings, then the results strengthen considerably ($\alpha_1 = 0.123$; $t = 2.05$; $\gamma_1 = -0.259$; $t = -3.71$), again supporting the notion that accruals are a performance control. Overall, the results in Table 11 indicate evidence of classification shifting, regardless of model used.

Results also support the existence of classification shifting when I examine those firms that just met the analyst forecast, especially for high market-to-book ratio firms (not tabulated). The interaction terms on *%SI* × *JUSTMET* and *%SI* × *JUSTMET* × *HIGH MB* are significantly positively associated with contemporaneous unexpected core earnings and significantly negatively associated with the unexpected change in core earnings in year $t+1$.[22] In sum, the models that do not include total accruals appear to do an inadequate job of controlling for performance, although some evidence of classification shifting remains. As discussed in Section I, since controlling for performance is a necessity of my model, I subject the results to additional subtests to corroborate my findings. However, at the end of the day, the reliance on an imperfect model is a limitation of my study.

---

[21] I roughly estimate accrual special items to be the sum of Compustat data items #213 (Loss [Gain] on sale of assets) and #217 (Funds from Operations, Other [which includes Reorganization Costs and Special Items]). This value is winsorized at net special items (Compustat data item #17) and set to zero if net special items are income-increasing.

[22] The total amount of initial shifting often remains negative in both settings. For example, when accruals are excluded from the model entirely, $\gamma_1 = -0.223 + \gamma_3 = 0.048$ for a net shifting of −0.175, as the main effect continues to be swamped by performance.

TABLE 11

Regression of Alternate Core Earnings and Future Change in Core Earnings Metrics on Special Items as a Percent of Sales, where Special Items are Classified as "Shiftable" or "Not Shiftable" Based on Type

Independent Variables	**Pred. Sign**	UE_CE_t No Accruals in the Model	UE_CE_t Accruals = $CE - CFO$	UE_CE_t Accruals = $NI - CFO - SI\_ACC$	UE_CE_t Accruals = $NI - CFO$
Intercept		0.037 (3.66)	0.026 (2.86)	0.023 (2.52)	0.025 (2.69)
$\%SI\_SHIFTABLE_t$	+	0.092 (1.44)	0.099 (1.69)	0.198 (3.38)	0.268 (4.43)
$\%SI\_NOT\_SHIFTABLE_t$	?	−0.293 (−4.02)	−0.338 (−5.05)	−0.036 (−0.55)	0.024 (0.34)
Adjusted R²		8.33%	13.06%	5.13%	8.52%

Independent Variables	**Pred. Sign**	$UE\_\Delta CE_{t+1}$ No Accruals in the Model	$UE\_\Delta CE_{t+1}$ Accruals = $CE - CFO$	$UE\_\Delta CE_{t+1}$ Accruals = $NI - CFO - SI\_ACC$	$UE\_\Delta CE_{t+1}$ Accruals = $NI - CFO$
Intercept		0.037 (3.17)	0.025 (2.25)	0.024 (2.87)	0.024 (2.47)
$\%SI\_SHIFTABLE_t$	−	−0.263 (−3.51)	−0.206 (−2.89)	−0.133 (−2.28)	−0.132 (−2.11)
$\%SI\_NOT\_SHIFTABLE_t$	?	0.276 (3.24)	0.252 (3.11)	0.218 (3.29)	0.181 (2.53)
Adjusted R²		10.99%	8.70%	7.69%	5.10%

The sample consists of 190 firm-year observations from 1996–2000, where the sample firms are in the S&P 500 in each corresponding year and had income-decreasing Compustat special items of at least 5 percent of sales in that year. $\%SI\_NOT\_SHIFTABLE_t$ represents special items reported by the firm that were either an asset write-off (excluding inventory and receivable write-offs) or a loss on the sale of an asset, scaled by sales in the same year. All other special items that are reported by the firm are classified as $\%SI\_SHIFTABLE_t$ and are also scaled by sales.

All variables are winsorized at 1 percent and 99 percent.

See Table 10 for a description of the alternate dependent variables.

VI. CONCLUSION

This paper examines the classification of items within the income statement as an earnings management tool. Unlike accrual management or the manipulation of real activities, classification shifting does not change bottom-line earnings and, thus, does not reverse in future periods or invite the same level of scrutiny by auditors and regulators. However, individual line items have different information content for future earnings and, correspondingly, for investors.

I examine classification shifting between core expenses (cost of goods sold and selling, general, and administrative expenses) and special items. Using a model of core earnings, analogous to that of the accrual model, I find that unexpected core earnings are increasing in special items in the year of the special item, and this unexpectedly high performance *reverses* in the following year. Furthermore, I find this shifting to be more pervasive when it allows the manager to meet the analyst forecast, as special items tend to be excluded from this earnings benchmark.

Overall, the evidence of classification shifting is compelling: (1) unexpected core earnings are increasing with special items in year t, but this improvement reverses in the following period; (2) the unexpected improvement only reverses if there are no special items present in year $t+1$, otherwise managers appear to classification shift again in year $t+1$, thereby maintaining the inflated core earnings; (3) these results hold only for those special items that, upon inspection, appear to be amenable to classification shifting; (4) these results are stronger for firms that just met the analyst forecast, and stronger yet for growth firms that just met the analyst forecast; and (5) there is some evidence that classification shifting is associated with negative returns in the subsequent year, suggesting that investors are negatively surprised when expenses that were previously excluded from core earnings recur.

There are many possible avenues for future research. First, the current paper relies on an imperfect model of core earnings; while this model is a first step to documenting classification shifting, improvements of this model might be examined in future studies. Future research might also further explore the negative abnormal returns documented herein by perhaps focusing on incentives to shift or by examining whether these returns vary cross-sectionally, for example with the sophistication of investors.

It is also possible that managers classification shift to a greater degree preceding equity issuances (Rangan 1998; Teoh et al. 1998) or when the valuation multiple on earnings is higher—such as for profit versus loss firms (Hayn 1995) or for those firms with high P/E ratios. In addition to capital market incentives, managers might wish to meet the analyst forecast in order to maximize their cash bonus (Matsunaga and Park 2001), sell their shares (Richardson et al. 2004; McVay et al. 2006), or maximize their compensation in general, as compensation committees tend to shield managers from income-decreasing transitory charges (Dechow et al. 1994; Gaver and Gaver 1998). These settings are left for future research.

The classification shifting documented in this paper is part of a much larger phenomenon of earnings management using income-statement classification (e.g., Barnea et al. 1976; Givoly et al. 1999; Davis 2002). There are many additional settings of classification shifting that future research might investigate. For example, managers might shift core expenses to expense classifications other than special items, such as R&D, which is valued differently than other core operating expenses (e.g., Lev and Sougiannis 1996; Aboody and Lev 2000), or non-operating expenses and discontinued operations, which are also excluded from analysts' forecasts (e.g., Abarbanell and Lehavy 2002).[23] Managers might also shift expenses across segments to hide abnormal profits (e.g., Berger and Hann 2003; Botosan and Stanford 2005) or allay investor concerns about certain segments. Each of these actions can mislead investors about the performance and proper valuation of a company.

REFERENCES

Abarbanell, J., and R. Lehavy. 2002. Differences in commercial database reported earnings: Implications for empirical research. Working paper, University of North Carolina.

Aboody, D., and B. Lev. 2000. Information asymmetry, R&D, and insider gains. *Journal of Finance* 55 (6): 2747–2766.

Anderson, M., R. Banker, and S. Janakiraman. 2003. Are selling, general, and administrative costs "sticky"? *Journal of Accounting Research* 41 (1): 47–63.

Baber, W., P. Fairfield, and J. Haggard. 1991. The effect of concern about reported income on discretionary spending decisions: The case of research and development. *The Accounting Review* 66 (4): 818–829.

Barnea, A., J. Ronen, and S. Sadan. 1976. Classificatory smoothing of income with extraordinary items. *The Accounting Review* 51 (1): 110–122.

Bartov, E., D. Givoly, and C. Hayn. 2002. The reward to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (2): 173–204.

———, and P. Mohanram. 2004. Private information, earnings manipulations, and executive stock-option exercises. *The Accounting Review* 79 (4): 889–920.

Berger, P., and R. Hann. 2003. Segment profitability and the proprietary costs of disclosure. Working paper, University of Chicago and University of Southern California.

Bernard, V., and J. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13 (4): 305–340.

Botosan, C., and M. Stanford. 2005. Managers' motives to withhold segment disclosures and the effect of SFAS No. 131 on analysts' information environment. *The Accounting Review* 80 (3): 751–771.

Bowen, R., A. Davis, and D. Matsumoto. 2005. Emphasis on pro forma versus GAAP earnings in quarterly press releases: Determinants, SEC intervention, and market reactions. *The Accounting Review* 80 (4): 1011–1038.

Bradshaw, M., and R. Sloan. 2002. GAAP versus the Street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41–66.

Brooks, L., and D. Buckmaster. 1976. Further evidence of the time series properties of accounting income. *Journal of Finance* 31 (December): 1359–1373.

Brown, L., and M. Caylor. 2005. A temporal analysis of quarterly earnings thresholds: Propensities and valuation consequences. *The Accounting Review* 80 (2): 423–440.

Bulkeley, W. 2002. Questioning the books: IBM annual report shows stronger core earnings. *Wall Street Journal* (March 12).

Burgstahler, D., J. Jiambalvo, and T. Shevlin. 2002. Do stock prices fully reflect the implications of special items for future earnings? *Journal of Accounting Research* 40 (3): 585–612.

Bushee, B. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* 73 (3): 305–333.

Cameron, R., and N. Gallery. 2001. Are abnormal items used to manage earnings? Working paper, Griffith University.

Carter, M. E. 2000. Operating performance following corporate restructurings. Working paper, Columbia University.

---

[23] For example, Nektar Therapeutics announced on March 1, 2005 that they would restate $27.6 million in expenditures from 2002–2004 that they had originally classified as R&D, but which should have been classified as Selling, General, and Administrative.

Davis, A. 2002. The value relevance of revenue for Internet firms: Does reporting grossed-up or barter revenue make a difference? *Journal of Accounting Research* 40 (2): 445–477.

DeAngelo, H., L. DeAngelo, and D. Skinner. 1994. Accounting choice in troubled companies. *Journal of Accounting and Economics* 17 (1-2): 113–143.

Dechow, P., and R. Sloan. 1991. Executive incentives and the horizon problem: An empirical investigation. *Journal of Accounting and Economics* 14 (1): 51–89.

———, M. Huson, and R. Sloan. 1994. The effect of restructuring charges on executives' cash compensation. *The Accounting Review* 69 (1): 138–156.

———, R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, and D. Skinner. 2000. Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons* 14 (2): 235–250.

———, S. Richardson, and I. Tuna. 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* 8 (2-3): 355–384.

Doyle, J., R. Lundholm, and M. Soliman. 2003. The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8 (2-3): 145–174.

———, and M. Soliman. 2005. Do managers define "Street" exclusions to meet or beat analyst forecasts? Working paper, University of Utah and Stanford University.

Dye, R. 2002. Classifications manipulation and Nash accounting standards. *Journal of Accounting Research* 40 (4): 1125–1162.

Elliott, J., and W. Shaw. 1988. Write-offs as accounting procedures to manage perceptions. *Journal of Accounting Research* 26 (Supplement): 91–119.

———, and J. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135–155.

Fairfield, P., R. Sweeney, and T. Yohn. 1996. Accounting classification and the predictive content of earnings. *The Accounting Review* 71 (3): 337–355.

———, and T. Yohn. 2001. Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies* 6 (4): 371–385.

Fama, E., and J. MacBeth. 1973. Risk, return and equilibrium—Empirical tests. *Journal of Political Economy* 81 (3): 607–641.

———, and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

———, and ———. 2000. Forecasting profitability and earnings. *Journal of Business* 73 (2): 161–175.

Francis, J., J. Hanna, and L. Vincent. 1996. Causes and effects of discretionary asset write-offs. *Journal of Accounting Research* 34 (Supplement): 117–134.

Freeman, R., J. Ohlson, and S. Penman. 1982. Book rate-of-return and prediction of earnings changes: An empirical investigation. *Journal of Accounting Research* 20 (2): 639–653.

Gaver, J., and K. Gaver. 1998. The relation between nonrecurring accounting transactions and CEO cash compensation. *The Accounting Review* 73 (2): 235–253.

Givoly, D., C. Hayn, and J. D'Souza. 1999. Measurement errors and information content of segment reporting. *Review of Accounting Studies* 4 (1): 15–43.

Gu, Z., and T. Chen. 2004. Analysts' treatment of nonrecurring items in street earnings. *Journal of Accounting and Economics* 38 (Conference Issue): 129–170.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125–153.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (1-3): 85–107.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Review* 40 (1): 105–134.

Hwang, S. 1994. Borden to reverse, reclassify 40% of 1992 charge. *Wall Street Journal* (March 22).

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Kasznik, R., and M. McNichols. 2002. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40 (3): 727–759.

Kinney, M., and R. Trezevant. 1997. The use of special items to manage earnings and perceptions. *The Journal of Financial Statement Analysis* (Fall): 45–53.

Lev, B., and T. Sougiannis. 1996. The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics* 21 (1): 107–138.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (Supplement): 37–64.

Lougee, B., and C. Marquardt. 2004. Earnings informativeness and strategic disclosure: An empirical examination of "pro forma" earnings. *The Accounting Review* 79 (3): 769–795.

Matsunaga, S., and C. Park. 2001. The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. *The Accounting Review* 76 (3): 313–332.

McNichols, M., and G. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

———. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4-5): 313–345.

McVay, S., V. Nagar, and V. Tang. 2006. Trading incentives to meet the analyst forecast. *Review of Accounting Studies*. Forthcoming.

Moehrle, S. 2002. Do firms use restructuring charge reversals to meet earnings targets? *The Accounting Review* 77 (2): 397–413.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* 6 (1): 109–154.

Penman, S., and X. Zhang. 2002. Modeling sustainable earnings and P/E ratios with financial statement analysis. Working paper, Columbia University and University of California.

Philbrick, D., and W. Ricks. 1991. Using Value Line and I/B/E/S analyst forecasts in accounting research. *Journal of Accounting Research* 29 (2): 397–417.

Phillips, J., M. Pincus, and S. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521.

Rangan, S. 1998. Earnings management and the performance of seasoned equity offerings. *Journal of Financial Economics* 50 (1): 101–122.

Revsine, L., D. Collins, and W. Johnson. 2005. *Financial Reporting and Analysis*. Upper Saddle River, NJ: Prentice Hall.

Richardson, S., S. H. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Schrand, C., and B. Walther. 2000. Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components. *The Accounting Review* 75 (2): 151–177.

Serwer, A. 2002. Kodak: In the noose. *Fortune Magazine* (February 4).

Sherer, P. 2000. AmeriServe examination finds financial woes. *Wall Street Journal* (July 3).

Shumway, T. 1997. The delisting bias in CRSP data. *Journal of Finance* 52 (1): 327–340.

Skinner, D., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (2-3): 289–312.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Soliman, M. 2004. Using industry-adjusted DuPont analysis to predict future profitability. Working paper, Stanford University.

Teoh, S. H., I. Welch, and T. Wong. 1998. Earnings management and the long-run underperformance of seasoned equity offerings. *Journal of Financial Economics* 50 (1): 63–99.

Weiss, I. 2001. Managerial responses to a transitory earnings shock: Strategic manipulation between the core and non-core components of earnings. Ph.D. Dissertation, University of Chicago.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 533–566

An Empirical Analysis of an Incentive Plan with Relative Performance Measures: Evidence from a Postal Service

Ella Mae Matsumura
University of Wisconsin–Madison

Jae Yong Shin
University of Wisconsin–Madison

ABSTRACT: Using 1997–1999 annual performance evaluation data of 214 postal stores in Korea, we find that financial performance improves following the implementation of an incentive plan that includes relative performance measures, and that under this incentive plan, the degree of common uncertainty is positively associated with store profitability. We also find evidence that the incentive effect of the plan is mitigated in stores at which the employees' perceived unfairness is likely to be high. Finally, we find that the net benefits of introducing the incentive plan may be conditional on the degree of common uncertainty, which appears to be inversely related to the perceived unfairness attributable to the introduction of the contract with relative performance measures.

Keywords: *relative performance measures; incentive plan; perceived unfairness; degree of common uncertainty.*

Data Availability: *Data in this study are derived from a proprietary source.*

I. INTRODUCTION

Relative performance evaluation (RPE) entails evaluating individual or organizational unit performance relative to the performance of others. Justification for RPE can be found in economics and behavioral theories. From an economics perspective, RPE provides benefits by deriving information about agents' efforts when the agents face common uncertainty, such as common market conditions. RPE then allows a principal to establish contracts that are Pareto superior to those based only on an individual agent's output

---

We are especially grateful to Joong-Bum Choi and Ki-Hwan Hong for providing us with institutional insights and helpful suggestions. We also thank Madhav Rajan (editor), two anonymous referees, Ana Albuquerque, Steve Hansen, Bill Lanen, participants at the 2004 American Accounting Association Management Accounting Section Research Conference, the 2005 Information, Markets, and Organizations Conference at the Harvard Business School, and workshop participants at the University of California, Riverside, University of Notre Dame, the University of Wisconsin–Madison, the University of Wisconsin–Milwaukee, and the University of Saskatchewan for their helpful comments. We are grateful to the participating company for allowing us access to the data.

Editor's note: This paper was accepted by Madhav Rajan.

Submitted June 2004
Accepted December 2005

(Holmstrom 1982). Behavioral theories highlight intrinsic (nonmonetary) influences on effort. Although RPE offers potential benefits, it may motivate reduced effort when workers perceive competition with co-workers as unfair (Frederickson 1992). For example, a worker may perceive competition with co-workers as unfair if co-workers have much greater ability or if the working environments present different challenges. RPE may also motivate undesirable behavior, such as sabotaging co-workers, colluding with co-workers, and shirking (Gibbons and Murphy 1990).

Our study uses data from the entire population of 214 post offices (hereafter, stores) managed by a postal service provider in Korea, which added relative performance measures[1] to its incentive plan in 1998. We examine whether (1) including relative performance measures in the incentive plan is positively associated with financial performance at the store level (the incentive effect), (2) the degree of common uncertainty is positively associated with store financial performance, (3) perceived unfairness due to introducing the incentive plan mitigates its performance impact, and (4) the degree of common uncertainty is negatively associated with likely perceived unfairness.

Our research site provides us with a powerful field-based setting to test the theory of RPE incentives. First, regulation precludes the monopolistic postal service provider from discontinuing operation in unprofitable areas, resulting in greater cross-sectional variation (as compared to unregulated industries) in terms of exogenous store characteristics that may influence store performance.[2] Second, to the extent that performance is largely driven by "uncontrollable" exogenous factors, we expect economic benefits from RPE because it helps insulate agents' compensation from uncontrollable "common" uncertainty.

Our results indicate that introduction of the research site's incentive plan that introduced new relative performance measures is positively associated with store-level financial performance and that under the incentive plan, a store's degree of common uncertainty is positively associated with profitability. Furthermore, we find evidence that the incentive plan's impact on financial performance is lower in stores with likely high perceived unfairness. Finally, we find some evidence that a store's degree of common uncertainty is negatively associated with the probability that the store has a high level of perceived unfairness.

This study has the following implications. First, it suggests that higher levels of common uncertainty with an RPE incentive plan can induce agents to exert more effort, leading to better financial performance. Second, prior empirical literature has been silent on the determinants of behavioral responses caused by an RPE incentive plan. Our study suggests that the degree of common uncertainty within a reference group can play a critical role in determining agents' perceived unfairness. A high degree of common uncertainty within each reference group appears to decrease the costs due to, for example, perceived unfairness of benchmarked targets among employees.

Our study contributes to the RPE literature by providing empirical evidence on RPE's performance impact when agents face varying degrees of common uncertainty. Furthermore, by examining whether RPE's incentive effect is moderated in stores where employees perceive unfairness, this study contributes to a growing literature that investigates the effect of compensation policies on performance and dysfunctional responses to compensation schemes. Our study also extends the literature on the importance of the degree of common uncertainty (Frederickson 1992) in determining the effectiveness of RPE contracts. In addition, this study contributes to the budgeting literature by providing some of the first empirical evidence on motivational effects of benchmarked accounting performance standards (Hansen et al. 2003). Finally, our study contributes to the growing literature that emphasizes the important role performance measurement systems can play in improving the efficiency of government operations. For example, Ittner and Larcker (1998a) suggest that the most fundamental question in governmental performance measurement practice is whether private sector notions of performance measurement and accountability are applicable in the public sector.

---

[1] We use "relative performance measures" to refer to measures that are used in comparing one agent or business unit to another, and "relative performance evaluation" to refer to a scheme that bases a person's compensation or nonmonetary performance evaluation at least in part on relative performance measures.

[2] A "cream-skimming" strategy, often observed in unregulated industries, will likely lead to a concentration of stores in profitable areas, resulting in less cross-sectional variation in terms of stores' external business environment.

The remainder of the paper is organized as follows. We review prior literature in Section II and describe our research site in Section III. Section IV develops our hypotheses. Section V describes our sample and variables, and presents descriptive statistics. Section VI presents our estimation models and results. Concluding remarks are offered in Section VII.

II. PRIOR LITERATURE

Due to data availability constraints, the extant archival empirical literature based on agency-theoretic predictions about RPE has largely focused on investigating whether top executives are compensated *as if* their performance is evaluated relative to competitors' performance and provides mixed results (Antle and Smith 1986; Gibbons and Murphy 1990; Janakiraman et al. 1992; Sloan 1993; Aggarwal and Samwick 1999; Anderson et al. 2001). For example, using data from 1947 to 1977 for 39 firms, Antle and Smith (1986) find weak evidence that the executives' compensation falls as other firms perform better, holding their own performance fixed. Using a comprehensive survey, Gibbons and Murphy (1990) find that executives are penalized when a competitor group performs better. However, Janakiraman et al. (1992) find little empirical evidence that the market and industry components of firm performance are completely removed in determining CEO compensation.

Research on the performance consequences of RPE or the relationship between RPE use and common uncertainty includes Casas-Arce and Martínez-Jerez (2005), who provide theoretical and empirical results on retailers compensated via dynamic tournaments, a form of RPE. Frederickson (1992) provides experimental evidence that agents' effort levels are higher with an RPE contract than with a profit-sharing contract where compensation is solely tied to agents' absolute performance. He also finds that subjects' effort levels increased significantly as the degree of common uncertainty increased with the RPE contract but not with the non-RPE contract. Using CEO compensation data, Kren (2002) finds that the degree of emphasis on RPE is positively related to the level of common uncertainty.

Combined, little empirical evidence is available on the performance consequences of RPE incentive plans when there is an important source of common uncertainty, especially at non-executive levels.[3] Moreover, to our knowledge, no study has empirically documented how dysfunctional responses offset an RPE contract's incentive effect. In addition, in their survey on budgeting practices, Hansen et al. (2003, 106) point out that even though relative performance measures as performance standards provide "an increased degree of legitimacy" due to the inherent credibility of performance achieved by similar units, there has been no accounting research that investigates the motivational impacts of benchmarked standards.

---

[3] Several studies that did not focus on RPE have found substantial incentive effects at non-executive levels. Banker et al. (1996) report that the implementation of a performance-based compensation plan increased sales and that the effect persisted and grew over time. Lazear (2000) finds that automobile worker output increased after the company switched its compensation scheme from hourly wages to piece rates. Banker et al. (2000) find that firm performance improved after customer satisfaction was included as a performance measure for managers. Sprinkle (2000) provides experimental evidence regarding how incentive-based compensation contracts compare to flat-wage compensation contracts in motivating individual performance. He shows that incentives not only increase the duration of effort, but also apparently increase the intensity of effort.

Our study fills these gaps in the extant RPE literature by using archival data to provide evidence on the performance impact of RPE given common uncertainty and by examining whether behavioral responses such as employees' perceived unfairness can mitigate the incentive effect. Our study also explicitly explores the degree of common uncertainty as a key factor influencing the net benefits of introducing relative performance measures into a contract. Recent research on RPE in executive compensation settings documents that previous mixed findings may be due to inappropriate choice of a peer group facing similar uncertainty (Albuquerque 2005). Our study supports the view that use of an appropriate peer group can significantly influence the motivational effectiveness of an RPE incentive plan.

III. RESEARCH SITE

Description of Research Site

The research site for this study is a postal service provider in Korea (hereafter, POST). POST is the state-owned organization that provides comprehensive postal services throughout Korea. POST is chartered as a monopolistic service provider to handle and deliver mail. However, it competes against several large domestic couriers and international couriers, such as Federal Express and DHL, in its expedited delivery market, which is more profitable than regular mail service. Intense competition in the expedited delivery market significantly undermines POST's profitability. Furthermore, rapid diffusion of electronic communication such as email and mobile phone service poses a serious threat to POST's core business. A universal service obligation imposed on POST precludes POST from using its discretion to close unprofitable stores, such as those located on remote islands. These factors, coupled with stagnant mail volume and decreased profit margins caused by strict regulation of postal service rates, contributed collectively to troubled financial performance.

POST's stores are homogeneous in many important operational aspects, such as organizational structure, mail handling and delivery process, incentive system, and basic infrastructure. However, stores vary significantly in terms of exogenous factors such as geographical location, clientele, and competition. Considering that mail volumes are largely determined by exogenous factors, it is highly likely that a store's profitability is conditional on such "uncontrollable" exogenous factors.[4]

Previous Incentive Plan

Even before it introduced the new RPE incentive plan in 1998, POST's incentive plan tied each employee's incentive pay to the performance of the store where the employee worked, and not to individual performance. All store-level front-line employees were covered by this incentive plan, as stipulated by the *Special Act on the Postal Business Operation.*

---

[4] We regressed each store's mail volume handled per employee on the following major exogenous factors: population, service area, average delivery distance, number of financial institutions (a proxy for income level), and location of store (large/middle/small city, and rural area) using 1997 data (prior to the new plan). The adjusted R^2 is 77 percent, suggesting that a substantial proportion of variation in mail volume can be attributed to exogenous factors over which a store manager cannot exercise control. Nevertheless, a store manager can influence costs and revenues.

Each store's performance score was a weighted average of multiple measures. One of the key performance measures for evaluating each store was its profitability growth rate relative to its previous-year profit, where store profitability was defined as store revenue divided by store operating cost. Profitability growth rate was selected because store profitability is influenced by "uncontrollable" exogenous factors and POST is obligated to provide universal service throughout the country. Therefore, absolute performance does not accurately reflect managerial effort, that is, "controllable" performance. Similarly, productivity growth rate, where store productivity was defined as mail volume handled per employee, was another key performance measure for each store.

At the end of each year, each store was ranked based on its weighted average evaluation score. Incentive pay was provided only for stores whose relative rank fell within the top 50 percent of all stores. Employees received the incentive bonus, if any, in April of the following year. If a store qualified for a bonus, then each employee in the store received the same percent of their monthly salary. Bonuses depended on the store's rank and ranged between 50–150 percent of monthly salary, which suggests that the maximum annual bonus constituted about 12.5 percent of annual compensation. Management concluded that the range of incentives provided had great potential to motivate workers, especially because performance-contingent bonuses were unusual among Korean state-owned companies.

By including the profitability growth rate relative to the previous year as a key performance measure, this incentive plan partially resolved the problem of comparing absolute performance of stores whose business environments are quite diverse. However, it ignored the relative efficiency among stores and accordingly did not effectively motivate stores to strive toward achieving greater efficiency. The following comments by a senior manager exemplify the motivational problems, which are similar to those that arise when a reward system ratchets performance standards, inducing employees to reduce effort and expected output in one period to avoid a higher standard in the next period (Indjejikian and Nanda 1999):

> Before we introduced the new incentive plan, individual stores opportunistically manipulated the growth rate in profitability ... Stores chose to exert effort on enhancing their profitability every two years. By doing so, in one year, they were able to maximize the growth rate in profitability relative to the previous year and get incentive pay. The next year they just didn't do anything ... Since all of us know that profit growth cannot be sustained eternally, basing incentive pay on profitability growth has an intrinsic problem.[5]

New Incentive Plan with Relative Performance Measures

In an attempt to resolve the motivational problems, POST introduced its new incentive plan for employees at each store in January 1998, with the objective of rewarding stores for their performance relative to an appropriate reference group.[6] Using cluster analysis, the new incentive plan classified all stores into nine reference groups, each of which shares a similar business environment.[7] Both relative profitability (store profitability divided by the average profitability in the reference group, where store profitability is defined as store revenues divided by store operating cost) and relative productivity (mail volume handled per store employee, divided by average productivity in the reference group) are key performance measures on which the largest weights are placed in determining overall performance tied to incentive pay. However, bonus eligibility remains the same as under the old plan. That is, incentive pay is provided only to the stores whose weighted averages are in the top 50 percent of all stores.

---

[5] To illustrate, a store's operating profit might be 100, 150, 100, and 150 in years 1990, 1991, 1992, and 1993, respectively. This pattern of profits is consistent with the following explanation. In 1991, the agents exerted high effort to maximize profit growth in order to get a bonus. In 1992, effort-averse agents did not exert any effort above the minimally required effort because they knew there was little room to improve the performance that year. In 1993, they exerted high effort again to maximize profit growth. This opportunistic behavior is attributable to the incentive plan that measures current year's performance relative only to the previous year's.

[6] The reference groups remained fixed until November 1999, when the groups were revised.

Table 1 summarizes the key similarities and differences between the new and old incentive plans. Both plans incorporate a tournament-type feature in which bonuses are paid only if stores rank in the top 50 percent on the weighted average evaluation score. However, the new incentive plan develops performance scores that specifically incorporate relative performance measures, that is, performance measures relative to stores with similar business environments. The new plan places 48 percent of the relative weight on a store's profitability and productivity measures relative to its reference group.[8] The relative weight placed on store growth rates over the previous year was reduced to 12 percent, compared to 60 percent under the old plan.

Figure 1 diagrams POST's schedule of introducing the new incentive plan. In December 1997, the framework was circulated to top management. Stores received details of the plan in January 1998. Therefore, in testing the impact of the new incentive plan, January 1998 is considered as the start of the incentive plan. Performance evaluation of stores in calendar year 1997 under the new plan was performed in March 1998 and the resulting bonuses were paid in April 1998.

IV. HYPOTHESES

H1 and H2: RPE Incentives and Degree of Common Uncertainty

To summarize our context, both the new and the old incentive schemes involve a tournament structure in which bonuses are awarded only to employees at stores ranked in the top 50 percent, based on the designated performance measure. The new scheme classifies the stores into nine reference groups, each of which shares a similar business environment. The new incentive plan additionally incorporates relative performance measures within each reference group in developing the performance scores used in the tournament ranking.

To develop our hypotheses, we begin by drawing on agency theory. Agency theory describes benefits associated with evaluating risk-averse agents on their relative performances when the agents' performances are affected by a common "shock" term (Holmstrom 1982). These benefits arise when information from a particular agent's output is supplemented with relative performance measures that the principal can use to reduce the agent's exposure to individual output uncertainty. Reducing the agent's compensation risk motivates the agent to exert higher effort, relative to the contract without the relative performance

---

[7] Group 1 contains stores in Seoul, which has approximately 11 million people. Groups 2 and 3 contain stores in suburbs of Seoul and the next three largest cities (2–4 million people in each city). Groups 4 through 6 contain stores in smaller, mid-sized cities, and Groups 7 through 9 contain stores in rural areas. The number of stores included in each reference group varies from 16 to 33. The arithmetic mean of profitability measures of each reference group varies from 0.4412 to 1.5754 (year 1997). Variables used in the cluster analysis include key factors that are expected to influence store profitability and productivity. Examples include population, square kilometers of the area covered by the store, and number of employees. Those variables are also used as control variables in our regression models in Section VI.

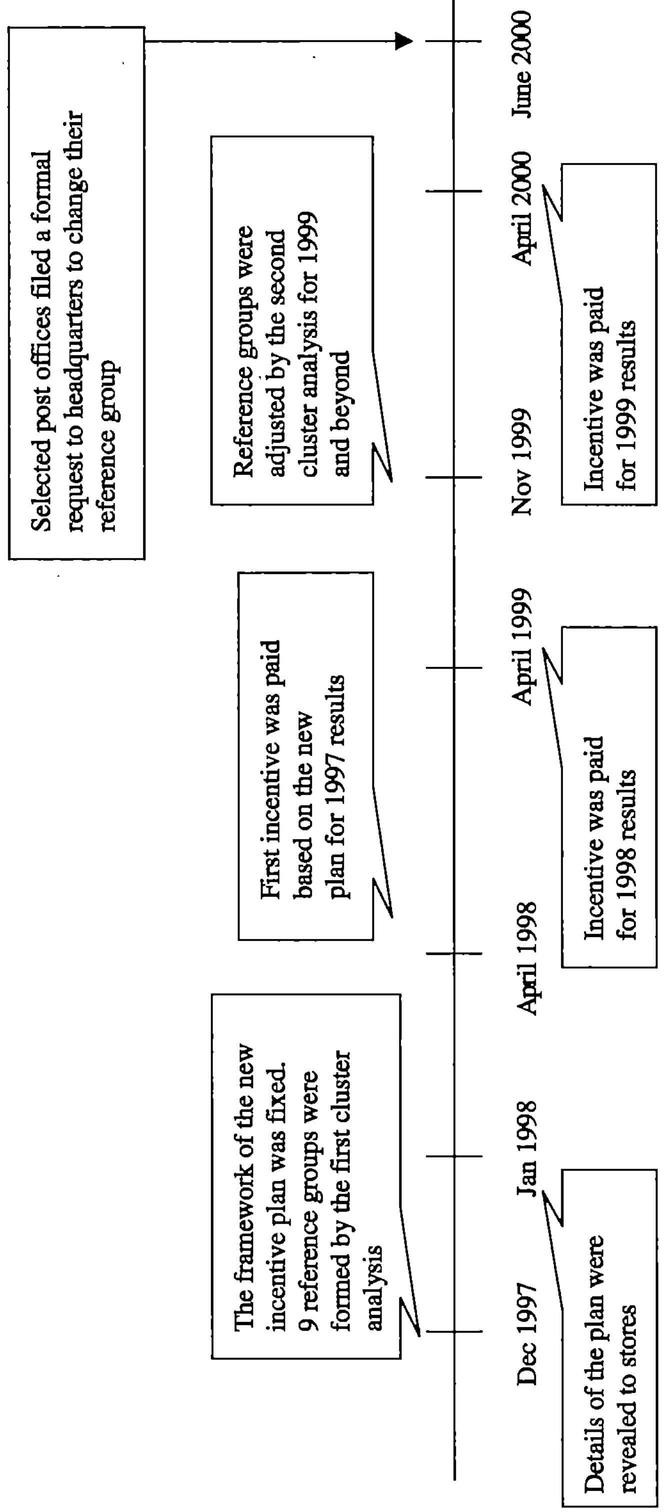
[8] In June 2000, the company allowed stores to request to change reference groups, especially if the stores reported a significant level of employees' perceived unfairness due to perceived unattainable performance standards.

TABLE 1
Comparison of New Incentive Plan with Old Plan

	Old Plan		New Plan	
Performance Dimension	**Performance Measure**	**Weights Placed on the Measure**	**Performance Measure**	**Weights Placed on the Measure**
Profitability	$\frac{\textit{Store Profitability } (t)}{\textit{Store Profitability } (t-1)}$	30%	$\frac{\textit{Store Profitability } (t)}{\textit{Reference Group Profitability } (t)}$	24%
			$\frac{\textit{Store Profitability } (t)}{\textit{Store Profitability } (t-1)}$	6%
Productivity	$\frac{\textit{Store Productivity } (t)}{\textit{Store Productivity } (t-1)}$	30%	$\frac{\textit{Store Productivity } (t)}{\textit{Reference Group Productivity } (t)}$	24%
			$\frac{\textit{Store Productivity } (t)}{\textit{Store Productivity } (t-1)}$	6%
Cash Flow Target	$\frac{\textit{Actual CFO } (t)}{\textit{Target CFO } (t)}$	20%	$\frac{\textit{Actual CFO } (t)}{\textit{Target CFO } (t)}$	20%
Nonfinancial	Customer Satisfaction	5%	Customer Satisfaction,	5%
	Employee Satisfaction	5%	Employee Satisfaction	5%
	Percentage of mail volume that reaches destination within specified time period	10%	Percentage of mail volume that reaches destination within specified time period	10%
Total		100%		100%
RPE component in performance measure?	No: each store's performance score is unaffected by performance in other stores		Yes: average performance of others in a store's reference group affects a store's performance score	
Rank store performance?	Yes		Yes	
Bonus eligibility	Top 50% of all stores		Top 50% of all stores	

Store Profitability (t) = store revenue divided by operating cost in year *t*;
Store Productivity (t) = mail volume handled per employee in year *t*;
Reference Group Profitability (t) = arithmetic mean of profitability of stores that belong to the specific reference group in year *t*;
Reference Group Productivity (t) = arithmetic mean of productivity of stores that belong to the specific reference group in year *t*; and
CFO (t) = cash flow from operations in year *t*.

FIGURE 1
Summary Schedule of New Incentive Plan Introduction


The framework of the new incentive plan was fixed. 9 reference groups were formed by the first cluster analysis
Details of the plan were revealed to stores
Dec 1997
Jan 1998
April 1998
First incentive was paid based on the new plan for 1997 results
Incentive was paid for 1998 results
April 1999
Reference groups were adjusted by the second cluster analysis for 1999 and beyond
Nov 1999
Incentive was paid for 1999 results
April 2000
Selected post offices filed a formal request to headquarters to change their reference group
June 2000

measures. Because POST's new incentive plan uses relative performance measures that the old plan did not, the new plan is expected to induce employees to provide greater effort, and hence expected improvement in financial performance.

The agency theory arguments above rest on the use of relative performance measures to reduce risk imposed on agents, thereby inducing higher effort. This perspective leads to a conclusion that there is no intrinsic value to inducing competition by linking each agent's reward to other agents' performance. However, behavioral theories argue that RPE can increase effort because agents interpret evaluation based on relative performance measures as a message that competing is appropriate behavior, again leading to higher effort and expected improvement in financial performance (Seta 1982; Frederickson 1992).

A further issue that arises in our context is the concern about the ratchet effect described in our earlier anecdote, which reports varying effort from year to year in order to increase the likelihood of receiving bonuses. A "reasonably objective" way to set targets or standards to encourage effort in multiperiod contexts that are susceptible to a ratchet effect is to use RPE (Milgrom and Roberts 1992, 233), as POST did by placing a relatively large weight on relative performance measures in determining which stores would receive bonuses. Thus, the economic and behavioral arguments all point to increased effort and expected improved financial performance with the new incentive plan, leading to Hypothesis 1.

> **H1:** *Ceteris paribus*, introduction of the new RPE incentive plan is positively associated with financial performance.

To develop Hypothesis 2, we first note from agency theory that if the agents' output measures are independent (that is, there is no common uncertainty), the principal does not gain from relative performance compensation schemes because no additional information is gained by comparing agents' outputs, precluding a risk-reducing and effort-increasing change in the contract (Holmstrom 1982). Conversely, Mookherjee (1984) develops conditions in a two-agent principal-agent model under which the principal's expected utility varies continuously with the correlation between the two agents' output measures. If the correlation is close to 1 (that is, there is maximal common uncertainty), then the principal can almost achieve the first-best welfare. Consistent with this theory, we expect that increases in the degree of common uncertainty are associated with increases in agents' efforts and, consequently, financial performance,

We next draw on referent cognitions theory, which states that a person's "negative reactions to a resource allocation decision occur when two conditions are met: (1) the outcomes associated with the decision are considerably lower than easily imagined alternative outcomes and (2) the procedures that give rise to outcomes are unfair, thereby rendering the outcomes unjustified" (Brockner and Wiesenfeld 1996, 193). In a budgeting or performance evaluation context, the first condition ("distributive justice") is related to fair budgets, standards, or targets, and the second condition ("procedural justice") is related to a fair budgeting or target-setting process. Based on referent cognitions theory, Libby (2001) predicts and finds in a budgeting setting that subjects' performance was lowest in the condition with an unfair target arising from an unfair process. We argue that in our situation, conversely, increases in the degree of common uncertainty increase fairness through both procedural and distributive justice. We therefore expect increases in the degree of common uncertainty to be associated with higher agent effort and, therefore, improved financial performance. Based on these behavioral arguments and our economic arguments, we predict that increases in the degree of common uncertainty are associated with increases in financial performance, as stated in Hypothesis 2 below:

H2: With the new RPE contract, *ceteris paribus*, increases in the degree of common uncertainty are positively associated with financial performance.

H3 and H4: RPE, Perceived Unfairness, and Degree of Common Uncertainty

Although economic theory shows that RPE can provide benefits when there is common uncertainty, several factors rest outside the theory. Most pertinent to our study and consistent with referent cognitions theory, RPE may induce reduced effort that, in turn, causes lower expected financial performance when workers perceive that the process of setting the relative performance targets and the targets themselves are unfair.[9] Drawing on referent cognitions theory, several studies on budgeting provide evidence that unfair budget targets are associated with lower performance when employees perceive the budgeting process is unfair (Magner et al. 1995; Libby 2001). Under the new incentive plan, stores in POST are assigned the average performance within a reference group as their budget target (that is, the threshold to achieve a bonus). Although this plan was intended to create fair targets, some stores viewed their targets as unfair due to a process with low procedural justice (for example, there was no store employee participation in the assignment of stores to reference groups). This reasoning suggests that performance will be lower in stores that perceive unfairness in the process and targets associated with the new incentive plan.

We apply the arguments above to develop Hypothesis 3 below. That is, employees' perceived unfairness is expected to be associated with lower financial performance. This prediction is also consistent with the findings that performance decreases at extremely difficult budget levels due to discouragement and less commitment to achieve the target (Merchant and Manzoni 1989; Fisher et al. 2003).

H3: The new RPE incentive plan's impact on performance is mitigated (lower) in stores at which the perceived unfairness is likely to be high, relative to other stores.

Referent cognitions theory stresses the importance of procedural justice in addition to distributive justice in relation to feelings of unfair treatment (Cropanzano and Folger 1989; McFarlin and Sweeney 1992). In our situation, all the stores are assigned into one of nine reference groups, and the perceived fairness of the evaluation procedure is likely to vary according to the extent to which a store is sorted into a more or less homogeneous reference group. Some stores reported major problems with dysfunctional responses due to decreased morale or skepticism, resulting from the perception that assignment to a particular reference group created unfair and unattainable standards for achieving a bonus. The perception of unfair assignment to a group and associated unattainable standards suggests that there was a low degree of common uncertainty within the reference group. Employees of stores assigned to a reference group with high common uncertainty are expected to view the process as more fair, in which case referent cognitions theory predicts less resentment and perceptions of unfairness, regardless of whether the target is ultimately achievable. We therefore predict that the degree of common uncertainty is negatively correlated with the extent to which employees perceive unfairness, as stated in Hypothesis 4:

---

[9] RPE contracts may also generate incentives to sabotage the measured performance of coworkers, to collude with coworkers and shirk, and to apply for jobs with inept coworkers (Gibbons and Murphy 1990). However, these dysfunctional responses are not very likely to occur in POST. Except for reference group 1 (stores in Seoul), each reference group comprises geographically dispersed stores. It is therefore not likely that a store manager could sabotage the measured performance of other stores in the same reference group or collude with them.

> **H4:** The degree of common uncertainty is negatively associated with likely perceived unfairness.

V. METHODS

Sample

Annual performance evaluation data were obtained for 1997 through 1999 for POST's 214 stores.[10] Financial and operational data for individual stores were collected from POST's performance measurement system. Control variables such as demographics and key statistics for individual stores were collected from the company's database that tracks information on individual stores' competitive environments. In addition, company documents and interviews with store managers provided qualitative data on their expectations and comments about the incentive plan.

Dependent Variables

For H1 through H3, we use *PROFRATE,* defined as revenue divided by operating cost of a store, as our dependent variable. We use this measure for the following reasons. First, this measure is actually used in the new incentive plan as a profitability measure. Second, by using *PROFRATE,* we are able to reduce the likelihood of heteroscedasticity that might occur when using alternative financial measures such as revenue or operating profit.[11] Third, by using this measure we can capture the effect of agents' efforts on both revenue increase and cost reduction.[12]

Discussion with POST's management revealed that stores initially responded to the new incentive plan by reducing operating cost rather than trying to increase revenue because mail volume, a store's major revenue driver, is largely determined by "uncontrollable" exogenous factors.[13] Furthermore, stores reportedly exerted effort to increase other revenue sources such as railway or airline ticketing services or e-post (an online shopping mall that sells specialty products peculiar to that region). Therefore, *PROFRATE* should effectively capture the plan's impact on efforts for reducing cost and increasing revenue at the store level. We also examine the impact of the new incentive plan on productivity by using *MVOL*, which is defined as the natural log of mail volume handled per employee in a store, as an alternative dependent variable.

---

[10] We were allowed to access the performance evaluation data for 2000. However, mergers among stores and changes in performance measures that occurred in 2000 preclude us from using the data for our analysis. Furthermore, in 2000, POST expanded the coverage of bonus eligibility to 80 percent of the stores from the previous 50 percent of the stores, for two reasons. First, a significant proportion of stores (47 stores, or 22 percent of the population) did not receive an incentive bonus for three consecutive years after the new incentive plan's introduction. In response to soaring complaints from these stores and management's concern about dysfunctional responses due to perceived unfairness of the benchmarked targets, POST decided to expand the bonus coverage. Second, although the new incentive plan significantly contributed to improving POST's financial performance, some critics argued that increased competition and too much emphasis on "profitable" mail service induced by the new plan compromised the fulfillment of an essential public duty (i.e., a service obligation that guarantees customers a universal standard of postal service at a single price regardless of where they live). To at least partially mitigate this concern, POST decided to expand the coverage.

[11] Heteroscedasticity is a common problem encountered when the estimation is based on cross-sectional data. One way to reduce this problem is to use a deflated dependent variable. For example, Banker et al. (2000) use normalized revenue and cost deflated by the number of available rooms as a proxy for the size of a hotel. In the "Robustness Analysis" subsection, we also report similar results using operating profit deflated by square feet of a store as a dependent variable (Banker et al. 1996).

[12] Postal rates did not change during 1997–1999 and we adjust cost data in 1998 and 1999 for inflation using the consumer price index (CPI) computed by the Bank of Korea.

[13] Examples of cost reduction efforts include eliminating unneeded employees, substituting part-timers for regular workers, encouraging early retirement of senior employees, and trimming unnecessary costs.

For H4, which examines the association between the degree of common uncertainty and employees' perceived unfairness, we use proxies for perceived unfairness as the dependent variable. These proxies are described in the following section.

Explanatory Variables

Plan Adoption Indicator Variables

Hypothesis 1 predicts that holding other factors fixed, introducing the new RPE incentive plan is positively associated with stores' financial performance. We test this hypothesis using dummy variables *FYR_PLAN* and *SYR_PLAN* to identify changes in profitability associated with implementation of the new incentive plan, after accounting for changes due to the control variables. *FYR_PLAN* is a dummy variable representing the first year following new incentive plan implementation, taking the value 1 for 1998, and 0 otherwise. *SYR_PLAN* is a dummy variable representing the second year following new incentive plan implementation, taking the value 1 for 1999, and 0 otherwise.

Degree of Common Uncertainty

Hypotheses 2 and 3 predict that two contextual factors—degree of common uncertainty and perceived unfairness—are likely to moderate the performance impact of the new RPE incentive plan. Specifically, H2 predicts that under the new contract, increasing the degree of common uncertainty increases financial performance and H3 predicts that the incentive plan's impact on performance is mitigated in stores at which the perceived unfairness is likely to be high, relative to other stores.

We operationalize the degree of common uncertainty (*COMMON*) using three empirical proxies: adjusted R^2, normalized residual dispersion, and average Euclidean distance of stores in a group from the reference group mean of performance control variables in Equation (1). Our first two empirical proxies are obtained by estimating the following regression equation for each reference group by year.[14]

$$\begin{aligned} PROFRATE_{i,t} = {} & \beta_0 + \beta_1 SERVICE\_AREA_{i,t} + \beta_2 POP_{i,t} + \beta_3 N\_BANK_{i,t} \\ & + \beta_4 N\_EMP_{i,t} + \beta_5 DIST_{i,t} + \beta_6 N\_DELPOINT_{i,t} \\ & + \beta_7 N\_COUNTER_{i,t} + \beta_8 N\_DELSEC_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{1}$$

where:

PROFRATE = profitability rate for store *i*, measured as revenue divided by operating cost at year *t*;

SERVICE_AREA = square kilometers (in hundreds) of the area that a store is responsible for at year *t*;

POP = the natural log of population (in thousands) of the area that a store is responsible for at year *t*;

---

[14] Unlike in Equations (2) and (3), discussed in Section VI, we could not use variable *METRO*, which takes on the values of 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively, in estimating reference-group-specific Equation (1) because, for example, group 1 consists of stores in Seoul only.

N_BANK = number of financial institutions (in hundreds) in the area that a store is responsible for at year t;
N_EMP = number of employees of a store at year t;
$DIST$ = average daily delivery distance (in hundred kilometers) per mail carrier of a store at year t;
$N\_DELPOINT$ = average number of daily delivery points (in hundreds) per mail carrier of a store at year t;
$N\_COUNTER$ = number of available counters (in hundreds) of a store at year t;
N_DELSEC = number of delivery districts (in hundreds) of a store at year t; and
t = 1997–1998.

Our first proxy, the adjusted R^2 obtained by running the above reference-group-specific regression, captures the extent to which the variation in profitability is explained by the variation in independent variables within each reference group. Considering that the independent variables are those used in the cluster analysis to form the nine reference groups, our adjusted R^2 measure should effectively capture the degree of common uncertainty.[15] Since the magnitude of R^2 can be noisy,[16] we use a ranking on adjusted R^2 (*RSQRANK*) that ranges from 1 to 9, with higher values indicating greater R^2.

Gu (2002) proposes using regression residual dispersions with proper control for scaling as an alternative to the regression R^2 and as a comparable measure of explanatory power in earnings' value-relevance studies. Gu (2002) points out that the residual variances are subject to the scale of the dependent variable, with high levels of the dependent variable generally associated with high residuals. Accordingly, we control for this scaling effect by dividing the estimated residual variance σ_ε^2 by the sample mean of our dependent variable *PROFRATE* in constructing our second proxy. To facilitate interpretation, we use a ranking on normalized residual variance (*RESIDRANK*) that ranges from 1 to 9, with higher values indicating smaller residual variance.

Our third proxy is the average of store-level Euclidean distance from the arithmetic mean of performance control variables of stores that belong to a specific reference group.[17] A rationale behind this measure is that a reference group whose stores are more closely populated along the within-reference-group mean of various controls is likely to exhibit more homogeneity among its members. We use a ranking on distance (*DISTRANK*) that ranges from 1 to 9, with higher values indicating smaller distance from the mean.

---

[15] This situation is analogous to the RPE literature on executive compensation documenting that the relevant peer group should be companies in the same industry rather than the entire stock market, in the sense that there should be more correlation in common shocks (i.e., greater common uncertainty) within the same industry (Prendergast 1999).

[16] For example, in 1997, the adjusted R^2 of the nine reference groups ranges from 0.159 to 0.863. For all of our proxies for *COMMON*, we report the results using the rank values of our three proxies for the degree of common uncertainty instead of the raw values in Section VI. We also discuss the results using raw values in the robustness analysis.

[17] Euclidean distance is the distance between two p-dimensional items and is widely used for cluster analysis (e.g., in the k-means method). Specifically, our measure is computed using the within-reference-group (RG) average,

$$d(Store_i, RGmean) = \sqrt{\sum_{k=1}^{5} (x_k - \bar{x}_k)^2}, \text{ where } x_1 = SERVICE\_AREA,\ x_2 = POP,\ x_3 = N\_BANK,\ x_4 = DIST,$$

and $x_5 = N\_EMP$. We first regress *PROFRATE* on the eight performance controls included in Equation (1) and retain five variables that are statistically significant predictors of *PROFRATE* and use them to compute Euclidean distance. All variables are standardized before calculating Euclidean distance because it is a convention to standardize variables before performing cluster analysis and our various performance controls used in Equation (1) are measured in different units (Johnson and Wichern 2002).

Perceived Unfairness

Another key issue is how to identify stores at which the level of perceived unfairness is likely to be high (*UNFAIR*) under the new incentive plan. Interviews with store managers revealed that the key negative response arising from the new plan was "abandonment." That is, employees who feel that their store reference group is too "strong" for them are not well motivated to exert effort to exceed the reference group's benchmark performance standard. In our analysis, we use two empirical proxies for *UNFAIR*. First, we use indicator variable *CHANGE*, which equals 1 for the 13 stores that filed a formal request to headquarters to change reference group in mid-2000.[18] These stores were allowed to file a formal complaint to postal headquarters only after a regional postmaster carefully investigated the store's reported problems. We infer that they are likely to have suffered from high levels of perceived unfairness due to feelings such as abandonment or skepticism of the new plan.

We also create an ordinal variable, *UNSATIS*, to measure potential perceived unfairness. An internal survey of the new incentive plan's impact revealed that employees became the most frustrated when they did not receive a bonus even though they improved performance relative to the previous year. This survey result is consistent with the prediction of referent cognitions theory that "Resentment is maximized when people believe they would have obtained better outcomes if the decision maker had used other procedures that should have been implemented" (Cropanzano and Folger 1989). Accordingly, we compared each store's rank on its profitability divided by the mean of its reference group (the profitability measure under the new plan) with the store's rank on profitability growth, and define *UNSATIS* as the decile ranking based on the difference, with higher values indicating higher perceived unfairness.[19] In sum, we expect a store to exhibit high perceived unfairness if that store's profitability relative to its peers is low even though it achieves high profitability growth.

Control Variables

POST's universal adoption of a new incentive plan across all the stores precludes us from using stores that did not change the incentive plan as a natural control sample. We therefore include a comprehensive set of proxy variables that likely control for differences in the business environments of stores in the estimation model, in order to isolate the incentive effect attributable to the new plan (Dopuch and Gupta 1997).[20] We first include the variables that were actually used in the cluster analysis to form POST's nine reference groups as performance controls, because POST believes those variables are highly correlated with a store's profitability. Variables such as *POP*, *SERVICE_AREA*, and *METRO* (which takes on the values of 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively), are also widely used as controls in the literature on analyzing the productivity in global postal services (e.g., Maruyama and Nakajima 2002,

---

[18] As described in Figure 1, POST adjusted the reference group for each store by performing a second cluster analysis in November 1999. Therefore, the stores whose reference group was changed in November 1999 could be considered those at which perceived unfairness is likely to be high. However, since new variables were added in the second cluster analysis, the change of reference group could be attributable to the addition of these new variables in the analysis. Therefore, we chose to use the 13 stores that filed a formal request to postal headquarters to change reference group in 2000 as stores with high "perceived unfairness."

[19] We do not have complete nonfinancial and cash flow target data to construct the new plan's total score. The two most highly weighted measures, relative profitability and relative productivity, are highly correlated, and so are their ranks (the correlation is 0.90, based on 1997–1999 data). Moreover, relative profitability is much more controllable than relative productivity. We therefore use the rank of relative profitability alone in defining *UNSATIS*.

[20] In estimating benchmark performance standards for public school expenditures, Dopuch and Gupta (1997) used proxy variables to reflect differences in socioeconomic factors that are used by the State of Missouri in setting a funding formula.

2003). However, among these variables, we do not include the number of employees of a store (*N_EMP*) as a control, in order to reflect store-level efforts to reduce the number of employees in response to the new incentive plan.

We additionally control for other factors that may influence store performance. Our *METRO* variable, which captures the economic viability of the area in which a store operates, also controls for the level of competition. This is because major competitors of POST, such as domestic and international couriers, operate mainly in profitable urban areas. We control for average labor cost per employee of a store (*SKILL*), in million Korean won, to capture the difference in the ability of the work force given that pay level increases with job tenure in POST. Finally, we control for the effect of product mix on store profitability with *PROD_MIX*, which is measured by the relative proportion of more profitable mail services (e.g., parcel, registered mail, and express mail) to total mail volume received by a store. If data on postmaster tenure were available, it could be included as an additional explanatory variable as a proxy for managerial expertise. Postmaster tenure data are not available, but because POST strictly enforces mandatory job rotation policy about every three years, we do not expect postmaster tenure to be highly correlated with store profitability.

Although we include many factors that may influence store-level performance, one should use caution in claiming that these control variables include all exogenous factors that affect a store's profitability. In addition, some control variables may suffer from measurement error. For example, *N_BANK* is included to control for the effect of average income on profitability. More appropriate proxies for average per capita income, such as income tax paid or average price of houses, were not available. Our concern about not including a more direct proxy for per capita income is alleviated by the fact that the whole nation's average per capita income in 1998 relative to 1997 decreased due to the macroeconomic impact of the nation's financial crisis in late 1997.[21] If income is a significant determinant of a store's profitability, then that should bias against our results. Furthermore, the rapid substitution of electronic communication such as email and mobile phone service for traditional surface mail in the late 1990s should also bias against finding our results.

Descriptive Statistics

Descriptive statistics of the variables used in the analyses appear in Panel A of Table 2.[22] For comparison, we provide descriptive statistics for performance under the old contract and the new RPE contract by year. Overall, the average financial performance of stores improved in the two years following introduction of the new plan. On average, operating profit increased from 60 million under the old plan to 556 million in the first year and to 916 million in the second year (1997-constant Korean won), due to increased revenues and reduced operating costs. One interesting finding is that the average operating cost of stores decreased substantially (by 11.1 percent) in 1998, the first year after introduction of the

---

[21] Since the Korean economic crisis in 1997 heavily affected the nation's economy, including mail volume, one might argue that improved performance in 1998 may be simply attributable to mail volume increases due to overall economic recovery from the crisis rather than incentive plan change. However, the negative economic impact of the crisis was the most conspicuous in 1998, not in 1997, because the crisis occurred in late November 1997. For example, real GDP growth rate in the Korean economy in 1998 was −6.7 percent compared to 6.8 percent in 1996 and 5.0 percent in 1997 (Crotty and Lee 2001). This suggests that any improvement in POST's store performance in 1998 is not likely to be the artifact of overall economic recovery from the crisis because 1998, the first year after the incentive plan change, was the year that was affected the most severely by the crisis. For a detailed chronological description of major events during the crisis, see Baek et al. (2004).

[22] Due to store-level missing data in some of our control variables, the final number of observations used in estimating our regression models is 619.

TABLE 2
Descriptive Statistics

Panel A: Descriptive Statistics for Variables Used in the Multivariate Analyses

					New Plan							
	Old Plan (1997)				1998				1999			
Variable	n	Mean	Std.	Median	n	Mean	Std.	Median	n	Mean	Std.	Median
REVENUE	214	3,833.57	4,488.26	2,259.28	214	3,912.12	4,288.73	2,325.62	214	4,245.45	4,559.84	2,558.63
COST	214	3,773.09	2,216.69	3,130.90	214	3,356.14	2,026.15	2,756.34	214	3,329.68	1,992.05	2,644.98
PROFIT	214	60.48	2,680.25	−579.83	214	555.98	2,640.81	−216.17	214	915.77	2,895.48	−77.60
PROFRATE	214	0.84	0.44	0.75	214	0.98	0.47	0.88	214	1.07	0.49	0.97
N_COUNTER	213	19.05	11.04	16.00	213	24.56	63.71	16.00	213	19.95	14.75	16.00
POP	214	220.71	190.72	144.78	212	222.03	193.99	142.10	211	223.04	193.97	138.08
SERVICE_AREA	213	467.54	380.04	463.00	214	466.89	374.35	459.50	213	470.72	376.94	465.00
N_DELPOINT	211	467.97	1,255.26	223.00	209	371.62	673.88	201.00	209	307.08	504.28	176.00
N_BANK	212	70.37	55.00	55.00	211	69.34	59.72	53.00	209	72.38	60.89	52.00
DIST	212	37.47	45.81	31.00	213	36.28	45.23	29.65	211	33.81	44.52	20.87
N_EMP	213	127.68	65.82	113.00	214	119.99	62.01	108.00	212	118.48	61.66	104.00
N_DELSEC	213	57.14	28.87	49.00	214	56.34	30.17	49.00	212	54.61	29.13	46.50
MVOL	213	329,890.27	217,644.70	271,205.00	214	398,843.67	242,365.98	328,747.00	213	425,274.02	241,440.69	359,345.00
SKILL	213	20.26	2.59	20.02	214	18.95	2.77	18.65	212	19.21	2.43	18.83
PROD_MIX	213	0.08	0.10	0.07	214	0.07	0.03	0.07	214	0.07	0.03	0.07
RSQ	214	0.46	0.26	0.39	214	0.53	0.20	0.56	214	0.40	0.35	0.47
RESID	214	0.03	0.02	0.03	214	0.02	0.01	0.02	214	0.03	0.01	0.03
EUCLID_DIST	214	1.69	0.64	1.52	214	1.49	0.54	1.45	214	1.34	0.45	1.20
CHANGE	214	0.06	0.24	0.00	214	0.06	0.24	0.00	214	0.06	0.24	0.00
UNSATIS	214	5.53	2.87	5.50	214	5.50	2.88	5.50	214	5.52	2.87	5.50

(*continued on next page*)

Panel B: Store Profitability Conditional on the Degree of Common Uncertainty (DCU) after New Incentive Plan

***RSQRANK$_{t-1}$* as a Proxy for Common Uncertainty**

	High DCU (n = 223)		Low DCU (n = 205)		t-test	Wilcoxon
	Mean	Median	Mean	Median	p-value	p-value
PROFRATE	1.096	0.965	0.945	0.876	0.00	0.01
ΔPROFRATE	0.121	0.115	0.105	0.105	0.08	0.14

***RESIDRANK$_{t-1}$* as a Proxy for Common Uncertainty**

	High DCU (n = 209)		Low DCU (n = 219)		t-test	Wilcoxon
	Mean	Median	Mean	Median	p-value	p-value
PROFRATE	1.244	1.252	0.793	0.707	0.00	0.00
ΔPROFRATE	0.123	0.125	0.103	0.097	0.04	0.00

***DISTRANK$_{t-1}$* as a Proxy for Common Uncertainty**

	High DCU (n = 229)		Low DCU (n = 199)		t-test	Wilcoxon
	Mean	Median	Mean	Median	p-value	p-value
PROFRATE	1.260	1.247	0.752	0.693	0.00	0.00
ΔPROFRATE	0.122	0.125	0.103	0.099	0.04	0.00

Panel C: Store Profitability Conditional on Perceived Unfairness (PU) after New Incentive Plan

***CHANGE* as a Proxy for Perceived Unfairness**

	High PU (CHANGE = 1; n = 26)		Low PU (CHANGE = 0; n = 402)		t-test	Wilcoxon
	Mean	Median	Mean	Median	p-value	p-value
PROFRATE	0.824	0.821	1.037	0.934	0.00	0.05
ΔPROFRATE	0.097	0.102	0.114	0.112	0.13	0.15

***UNSATIS$_{t-1}$* as a Proxy for Perceived Unfairness**

	High PU (n = 215)		Low PU (n = 213)		t-test	Wilcoxon
	Mean	Median	Mean	Median	p-value	p-value
PROFRATE	0.825	0.684	1.224	1.178	0.00	0.00
ΔPROFRATE	0.082	0.072	0.115	0.118	0.00	0.00

(continued on next page)

TABLE 2 (Continued)

---

[a] Cost data in 1998 and 1999 are adjusted for inflation using the consumer price index (CPI), whereas revenue data in 1998 and 1999 are unadjusted because postal rates did not change during 1997–1999.

Variable Definitions:

REVENUE = revenue of a store (million Korean won);
COST = operating cost of a store (million Korean won);
PROFIT = operating profit of a store (million Korean won);
PROFRATE = rate of profitability for a store, measured as revenue divided by operating cost;
ΔPROFRATE = change in profitability;
SERVICE_AREA = square kilometers of the area that a store is responsible for;
POP = population of the area that a store is responsible for (thousands);
N_BANK = the number of financial institutions of the area that a store is responsible for;
N_EMP = the number of employees of a store;
DIST = average daily delivery distance per mail carrier of a store (kilometers);
N_DELPOINT = average number of daily delivery points per mail carrier of a store;
N_COUNTER = the number of available counters of a store;
N_DELSEC = the number of delivery sections of a store;
MVOL = mail volume handled per employee in a store;
SKILL = average labor cost per employee of a store (million Korean won);
PROD_MIX = the relative proportion of parcel, registered mail, and express mail to total mail volume received by a store;
RSQ = the adjusted R^2 of a store obtained by running reference group-specific regression Equation (1);
RESID = residual dispersion of a store obtained by running reference group-specific regression Equation (1);
EUCLID_DIST = the average Euclidean distance of stores in a group from the reference group mean of performance control variables;
CHANGE = dummy variable representing stores that requested to change their reference group in 2000, taking the value 1 if the store requested a change, and 0 otherwise; and
UNSATIS = the decile ranking of the difference between a store's rank on profitability growth and on relative profitability, with higher values indicating stores that exhibited high profitability growth, but low relative profitability.

new plan, consistent with our discussions with management that stores initially responded to the new incentive plan by reducing operating cost rather than by increasing revenue. However, this sharp decrease in operating cost relative to the previous year almost disappears in the second year (1999) and continued profit growth in the second year appears to be achieved through increased revenue rather than reduced cost.

The large standard deviation in operating profit implies that stores vary significantly in terms of profitability. The profitability rate (revenue divided by operating cost) increased from 0.84 during the under the old plan to 0.98 in the first year and to 1.07 in the second year following introduction of the new plan. An average profitability measure that is smaller than 1 in 1997 suggests that on average, stores were suffering losses. The results of the paired t-test, not tabulated, show that most of the control variables are not statistically different under the old and new plans. This provides preliminary evidence that the change in profitability, if any, could be attributable to the implementation of the new RPE plan.

Panels B and C of Table 2 present store profitability (*PROFRATE*) and change in profitability (Δ*PROFRATE*) by classifying the stores based on the degree of common uncertainty and perceived unfairness using the data after introduction of the new plan.

As seen in Panel B of Table 2, the high degree of common uncertainty (DCU) stores exhibit a statistically higher level of *PROFRATE* and Δ*PROFRATE* than the low DCU stores in the years following the new plan. Similarly, Panel C shows that the high perceived unfairness (PU) stores exhibit statistically lower levels of *PROFRATE* and Δ*PROFRATE*. These results are robust to the alternative measures of the degree of common uncertainty and perceived unfairness. Consistent with our H2 and H3, these findings provide preliminary evidence on the role of the degree of common uncertainty and perceived unfairness as moderating factors of performance improvement in response to the new incentive plan. These results, however, should be interpreted with caution because the analysis in Table 2 does not control for store attributes that potentially influence profitability.

Univariate Correlations

Table 3 provides details on Pearson correlations among the variables used in the analyses. Square kilometers of the area (*SERVICE_AREA*), average number of daily delivery points per mail carrier (*N_DELPOINT*), and average daily distance per mail carrier (*DIST*) are negatively and significantly correlated with profitability. Other independent variables such as number of financial institutions (*N_BANK*) and population (*POP*) are all positively and significantly correlated with a store's profitability.

As expected, correlations among alternative definitions of the degree of common uncertainty ($RSQRANK_{t-1}$, $RESIDRANK_{t-1}$, $DISTRANK_{t-1}$) are significantly positive and store profitability (*PROFRATE*) is positively correlated with all three variables. *PROFRATE* is also significantly negatively correlated with alternative definitions of perceived unfairness (*CHANGE* and $UNSATIS_{t-1}$). The correlation between *CHANGE* and $UNSATIS_{t-1}$ is positive but not significant.[23] $UNSATIS_{t-1}$ is significantly negatively correlated with all three proxies of the degree of common uncertainty, providing preliminary support for our H4. The correlation between *CHANGE* and our three common uncertainty variables is negative but not significant.

---

[23] Since we provide correlations among variables used in our Equations (2) and (3) in Table 3, we were unable to include correlations among measures of common uncertainty and perceived unfairness during the 1998–1999 period that is actually used for testing H4. In this time period, the correlation between *CHANGE* and *UNSATIS* is positive and significant (0.14, p-value = 0.004). *UNSATIS* is significantly negatively correlated with all three proxies of the degree of common uncertainty. The correlation between *CHANGE* and the three common uncertainty variables is negative, but it is significant only with *DISTRANK*.

TABLE 3
Pearson Correlations among Variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
1. *PROFRATE*	1.00																
2. *N_COUNTER*	0.23	1.00															
3. *POP*	0.73	0.23	1.00														
4. *SERVICE_AREA*	−0.56	−0.09**	−0.49	1.00													
5. *N_DELPOINT*	−0.09**	−0.04^{NS}	−0.07*	0.09**	1.00												
6. *N_BANK*	0.72	0.23	0.65	−0.31	−0.04^{NS}	1.00											
7. *DIST*	−0.44	−0.05^{NS}	−0.38	0.43	0.36	−0.27	1.00										
8. *N_EMP*	0.67	0.33	0.68	−0.24	−0.05^{NS}	0.80	−0.25	1.00									
9. *N_DELSEC*	0.63	0.32	0.74	−0.21	−0.05^{NS}	0.76	−0.22	0.91	1.00								
10. *MVOL*	0.94	0.25	0.70	−0.61	−0.10**	0.70	−0.45	0.67	0.60	1.00							
11. *SKILL*	0.31	0.13	0.32	−0.25	0.02^{NS}	0.34	−0.17	0.43	0.34	0.41	1.00						
12. *PROD_MIX*	0.11	−0.00^{NS}	0.09**	−0.03^{NS}	0.03^{NS}	0.07*	−0.09**	0.05^{NS}	0.06^{NS}	0.06^{NS}	0.01^{NS}	1.00					
13. *RSQRANK*$_{t-1}$	0.15	0.06^{NS}	0.32	−0.12**	−0.05^{NS}	0.23	−0.12**	0.28	0.22	0.22	0.25	0.06^{NS}	1.00				
14. *RESIDRANK*$_{t-1}$	0.59	0.20	0.61	−0.45	−0.18	0.50	−0.31	0.53	0.51	0.61	0.32	−0.14	0.30	1.00			
15. *DISTRANK*$_{t-1}$	0.50	0.11**	0.41	−0.49	−0.05^{NS}	0.35	−0.34	0.33	0.33	0.50	0.26	0.04^{NS}	0.17	0.36	1.00		
16. *CHANGE*	−0.10	−0.03^{NS}	−0.09**	−0.04^{NS}	0.02^{NS}	−0.07*	−0.01^{NS}	−0.05^{NS}	−0.07*	−0.07*	0.10**	−0.07*	−0.02^{NS}	−0.08^{NS}	−0.00^{NS}	1.00	
17. *UNSATIS*$_{t-1}$	−0.51	−0.03^{NS}	−0.32	0.25	0.00^{NS}	−0.34	0.23	−0.28	−0.21	−0.48	−0.20	−0.21	−0.16	−0.20	−0.09**	0.05^{NS}	1.00

*, ** Indicates that the correlation coefficient is significantly different from zero at the 10 percent and 5 percent levels, respectively (two-tailed).

^{NS} Indicates that the correlation coefficient is not significantly different from zero at the 10 percent level (two-tailed).

All other correlation coefficients are significantly different from zero at the 1 percent level (two-tailed).

Variable Definitions:

RSQRANK = the adjusted R^2 ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating greater R^2;

RESIDRANK = residual dispersion ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating smaller residual dispersion; and

DISTRANK = the average Euclidean distance of stores in a group from the reference group mean of performance control variables; ranges from 1 through 9 with higher values indicating smaller distance from the mean

For definition of other variables, see Table 2.

VI. EMPIRICAL MODELS AND RESULTS

Performance Impact of New Incentives

To test H1, we estimate the following fixed-effects regression equation using pooled cross-sectional time-series data for the 214 stores:

$$\begin{aligned} PROFRATE_{i,t} = {} & \alpha_0 + \alpha_1 FYR\_PLAN_{i,t} + \alpha_2 SYR\_PLAN_{i,t} + \alpha_3 SERVICE\_AREA_{i,t} \\ & + \alpha_4 POP_{i,t} + \alpha_5 N\_BANK_{i,t} + \alpha_6 METRO_{i,t} + \alpha_7 DIST_{i,t} \\ & + \alpha_8 N\_DELPOINT_{i,t} + \alpha_9 N\_COUNTER_{i,t} + \alpha_{10} N\_DELSEC_{i,t} \\ & + \alpha_{11} SKILL_{i,t} + \alpha_{12} PROD\_MIX_{i,t} + \varepsilon_{i,t} \qquad (2) \end{aligned}$$

where $t = 1997\text{–}1999$.

To control for any store-specific omitted variables, we estimate Equation (2) using fixed-effects regression. The coefficient on *FYR_PLAN* captures the average difference in profitability measures between the first year after the new incentive plan (year 1998) and base year 1997. Similarly, the coefficient on *SYR_PLAN* captures the average difference in profitability measures between the second year after the new incentive plan (year 1999) and base year 1997. Significant positive coefficients on *FYR_PLAN* and *SYR_PLAN* after controlling for other exogenous factors are interpreted as their average impact on financial performance, thus providing support for H1.

Moderating Effects of Degree of Common Uncertainty and Perceived Unfairness

To jointly test H2 and H3, we estimate the following fixed-effects regression equations using store-level data for 1998–1999:

$$\begin{aligned} PROFRATE_{i,t} = {} & \delta_0 + \delta_1 COMMON_{i,t-1} + \delta_2 UNFAIR_{i,t} + \delta_3 SERVICE\_AREA_{i,t} \\ & + \delta_4 POP_{i,t} + \delta_5 N\_BANK_{i,t} + \delta_6 METRO_{i,t} + \delta_7 DIST_{i,t} \\ & + \delta_8 N\_DELPOINT_{i,t} + \delta_9 N\_COUNTER_{i,t} + \delta_{10} N\_DELSEC_{i,t} \\ & + \delta_{11} SKILL_{i,t} + \delta_{12} PROD\_MIX_{i,t} + \varepsilon_{i,t} \qquad (3) \end{aligned}$$

where $t = 1998\text{–}1999$.

We use *RSQRANK*, *RESIDRANK*, and *DISTRANK* as alternatives for *COMMON*, and we use *CHANGE* and *UNSATIS* as alternatives for *UNFAIR*. We expect the coefficient on *COMMON* to be significantly positive, which suggests that with the new RPE contract, a store's degree of common uncertainty within its reference group is positively associated with its financial performance. Since the profitability of a store relative to its reference group receives the largest weight in determining bonuses under the RPE-based plan, we expect the coefficient on *UNFAIR* to be significantly negative. This would suggest that the new incentive plan's performance impact is mitigated in stores that suffered from a high level of perceived unfairness, indicating that the new RPE-based plan's incentive effect is conditional on the employees' perception of fairness of the benchmarked target.[24]

---

[24] Recall that H4 predicts a negative relationship between the degree of common uncertainty and perceived unfairness, both of which are included as key explanatory variables in Equation (3). It follows that common uncertainty may have both direct and indirect effects on store profitability. For example, a high degree of common uncertainty may positively affect store profitability directly, and also positively affect store profitability indirectly through lowered perceived unfairness. Our multiple linear regression model does not allow us to reliably disentangle the direct effect from the indirect effect of the degree of common uncertainty on performance. Accordingly, the effect of the degree of common uncertainty on performance as captured by the coefficient on *COMMON* may be understated because we explicitly control for perceived unfairness in the regression.

We use $COMMON_{t-1}$ and $UNSATIS_{t-1}$ in estimating Equation (3) under the assumption that the employees will exert effort based on their perception of the benchmarked performance standards in the previous year. Since actual performance evaluation of stores for year t is performed in March of year $t+1$, stores do not know the benchmarked performance standards against which their performance will be compared during year t. Thus, we assume that the benchmarked performance standard used in the year $t-1$ evaluation serves as a reasonable estimate of a performance standard for year t evaluation.

Association between Perceived Unfairness of Benchmarked Target and Degree of Common Uncertainty

To test the negative association predicted in H4, we estimate two regression equations using store-level data for 1998–1999. Unlike in Equations (2) and (3), in which store performance (*PROFRATE*) is our dependent variable, our dependent variables here are our two proxies for store-level perceived unfairness, namely, *CHANGE* and *UNSATIS*. We first estimate Equation (4), a binary logit regression examining the association between *CHANGE* and *COMMON*. We then estimate Equation (5), an ordered logit regression examining the association between *UNSATIS* and *COMMON*. Hypothesis 4 predicts that the coefficient on *COMMON* will be significantly negative, suggesting that agents are more (less) likely to suffer from perceived unfairness under the new RPE incentive contract if the degree of common uncertainty is low (high) within their reference group.[25]

$$\begin{aligned}\text{Prob}(CHANGE = 1) = F(&\alpha_0 + \alpha_1 COMMON_{i,t} + \alpha_2 SERVICE\_AREA_{i,t}\\ &+ \alpha_3 POP_{i,t} + \alpha_4 N\_BANK_{i,t} + \alpha_5 METRO_{i,t}\\ &+ \alpha_6 N\_EMP_{i,t} + \alpha_7 DIST_{i,t} + \alpha_8 N\_DELPOINT_{i,t}\\ &+ \alpha_9 N\_COUNTER_{i,t} + \alpha_{10} N\_DELSEC_{i,t}\\ &+ \sum_{i=1}^{7} \alpha_{i+10} REGION_i) \qquad (4)\end{aligned}$$

$$\begin{aligned}\text{Prob}(UNSATIS) = F(&\delta_0 + \delta_1 COMMON_{i,t} + \delta_2 SERVICE\_AREA_{i,t} + \delta_3 POP_{i,t}\\ &+ \delta_4 N\_BANK_{i,t} + \delta_5 METRO_{i,t} + \delta_6 N\_EMP_{i,t}\\ &+ \delta_7 DIST_{i,t} + \delta_8 N\_DELPOINT_{i,t} + \delta_9 N\_COUNTER_{i,t}\\ &+ \delta_{10} N\_DELSEC_{i,t} + \sum_{i=1}^{7} \delta_{i+10} REGION_i) \qquad (5)\end{aligned}$$

where t = 1998–1999.

One potential issue that arises is controlling for factors related to store-level variation in the budgeting process that may influence the perceived fairness of the budget target.

---

[25] We estimate logistic regressions of Equations (4) and (5) by including seven region dummies to control for unspecified geographical region-specific differences in stores' profitability rather than store-specific fixed effects because in a fixed-effects logit model, the ability to estimate parameters is corrupted by an inability to estimate consistently the subject-specific effects (Frees 2004). The seven regions, which roughly correspond to the provinces, are defined by the area covered by each of POST's regional headquarters. The *REGION* variables correspond to the eight regions, minus one that is the base group (the province that includes Seoul, the capital and largest city in Korea). These dummies are included to control for the effect of macroeconomic factors that are peculiar to the region in which a store is located.

However, because all the stores were required to adopt the same new plan, we do not have such store-level budgeting factors. We do attempt to control for potential effects of the different business environments facing each store—although exploratory, we include the nine exogenous variables used by POST in forming reference groups.

Impact of Incentive Plan Change on Financial Performance

Column (1) of Table 4 presents the estimation results of Equation (2) examining the new incentive plan's impact on profitability (that is, the profitability rate). The coefficients on *FYR_PLAN* and *SYR_PLAN*, after controlling for other exogenous factors, are positive and significant ($p < 0.0001$ for both, two-tailed), indicating that the incentive plan change has a positive and significant effect on financial performance. Thus, our results support H1.[26] To assess the economic significance of the result, the coefficients on *FYR_PLAN* and *SYR_PLAN* suggest that the introduction of the new incentive plan is associated with the increase in average store profitability by 0.137 in the first year (1998) and by 0.221 in the second year (1999) compared to the base year (1997) after controlling for other factors that may influence store profitability.

Column (2) of Table 4 shows that the results are very similar when mail volume handled per employee (*MVOL*) is used as an alternative dependent variable. This suggests that improved productivity caused by the new incentive plan plays a role in driving profitability.[27] As discussed in Section V, interviews with management and store managers revealed that the initial profitability improvement in response to the new incentive plan is mainly attributable to reducing labor and overhead cost. Given that mail volume is largely determined by exogenous factors, a positive and significant incentive effect on productivity appears to be driven by stores' efforts to eliminate unnecessary employee hours.

To provide insight into whether the new plan's impact on performance persisted and grew over time (i.e., continuing improvement), we test equality of the coefficients on *FYR_PLAN* and *SYR_PLAN*. Based on the result of the F-test, we reject the null hypothesis that the two coefficients are equal at the 1 percent level, suggesting that store employees exerted more effort to learn how to perform better and were motivated to further improve profitability in the second year following implementation of the new plan.[28]

Effect of Degree of Common Uncertainty and Perceived Unfairness on Financial Performance

Table 5 provides the results from estimating Equation (3) with the six combinations arising from crossing three measures of *COMMON* with two measures of *UNFAIR*. The coefficient on *COMMON* is positive and significant at the 5 percent level (two-tailed) except in column (1), where the coefficient is significant at the 10 percent level, and column (5), where the coefficient is not significant. These results suggest that holding other factors fixed, a greater degree of common uncertainty leads to higher store performance under the

---

[26] We compute variance inflation factors (VIF) and the condition index to investigate potential inference problems due to collinearity. Most of the variance inflation factors for the variables are less than 2 and the largest variance inflation factor is 5.1, which is smaller than the threshold of 10 in the econometrics literature (Kennedy 1992). The largest condition index number is 33, which is slightly larger than 30, the cut-off proposed by Belsley et al. (1980). We investigate the proportion of variation contributed by each variable and find that *SKILL* is responsible for the condition index being over 30. However, our findings are robust to excluding *SKILL* from the model.

[27] Since the use of mail volume handled per employee (*MVOL*) as an alternative dependent variable produces very similar results, the subsequent analyses use only *PROFRATE* as the dependent variable.

[28] When *MVOL* is used as the dependent variable, the coefficient on *SYR_PLAN* is still statistically larger than that on *FYR_PLAN*, but the difference between the two becomes smaller than when we use *PROFRATE* as our dependent variable. This suggests that the new plan's impact on continuing improvement of profitability in the second year was greater than its impact on productivity because, in comparison to the productivity measure over which stores could not exert much control, the profitability rate provides us with a more comprehensive performance measure in the sense that it reflects stores' efforts to trim overhead costs and to increase revenue sources other than those generated by mail volume.

TABLE 4
Regressions Examining the Impact of Incentive Plan Change on Performance
(t = 1997–1999)

$$\begin{aligned} PERF_{i,t} = {} & \alpha_0 + \alpha_1 FYR\_PLAN_{i,t} + \alpha_2 SYR\_PLAN_{i,t} + \alpha_3 SERVICE\_AREA_{i,t} + \alpha_4 POP_{i,t} \\ & + \alpha_5 N\_BANK_{i,t} + \alpha_6 METRO_{i,t} + \alpha_7 DIST_{i,t} + \alpha_8 N\_DELPOINT_{i,t} \\ & + \alpha_9 N\_COUNTER_{i,t} + \alpha_{10} N\_DELSEC_{i,t} + \alpha_{11} SKILL_{i,t} + \alpha_{12} PROD\_MIX_{i,t} \\ & + \varepsilon_{i,t} \end{aligned}$$

		(1)	(2)
		PERF	
		PROFRATE	*MVOL*
Variable	**Predicted Sign**	**Coefficient (t-statistic)**	**Coefficient (t-statistic)**
Constant	?	2.281* (1.84)	1.958 (1.38)
FYR_PLAN	(+)	0.137*** (6.17)	0.244*** (9.64)
SYR_PLAN	(+)	0.221*** (10.07)	0.304*** (12.12)
SERVICE_AREA	?	−0.017*** (−5.24)	−0.021*** (−5.55)
POP	?	0.093*** (5.15)	0.258*** (12.53)
N_BANK	?	0.239*** (9.90)	0.249*** (9.01)
METRO	?	0.194*** (9.95)	0.212*** (9.55)
DIST	?	−0.058** (−2.39)	−0.091*** (−3.28)
N_DELPOINT	?	0.0002 (0.18)	0.001 (0.81)
N_COUNTER	?	0.020 (0.84)	0.048* (1.77)
N_DELSEC	?	−0.026 (−0.47)	−0.497*** (−7.94)
SKILL	?	−0.179** (−2.41)	0.433*** (5.09)
PROD_MIX	?	0.215 (1.51)	−0.019 (−0.12)
n		619	619
Adjusted R^2		80.2%	83.8%
F-test of H_0: $\alpha_1 = \alpha_2$		$p < 0.0001$	$p = 0.004$
Change in R^2 by *FYR_PLAN* and *SYR_PLAN*		3.4%***	4.3%***

(*continued on next page*)

TABLE 4 (Continued)

---

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively (two-tailed).

Reported t-statistics are based on the White (1980) standard errors.

Variable *PERF* represents the rate of profitability (*PROFRATE*) or mail volume handled per employee (*MVOL*).

Variable Definitions:

FYR_PLAN = dummy variable representing the first year following RPE incentive plan implementation, taking the value 1 for 1998, and 0 otherwise;
SYR_PLAN = dummy variable representing the second year following RPE incentive plan implementation, taking the value 1 for 1999, and 0 otherwise;
SERVICE_AREA = square kilometers of the area that a store is responsible for at year *t* (in hundreds);
POP = the natural log of population (in thousands) of the area that a store is responsible for at year *t*;
N_BANK = number of financial institutions in the area that a store is responsible for at year *t* (in hundreds);
METRO = variable taking on the values of 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively;
DIST = average daily delivery distance per mail carrier of a store at year *t* (in hundred kilometers);
N_DELPOINT = average number of daily delivery points per mail carrier of a store at year *t* (in hundreds);
N_COUNTER = number of available counters of a store at year *t* (in hundreds);
N_DELSEC = number of delivery districts of a store at year *t* (in hundreds);
SKILL = average labor cost per employee of a store (million Korean won); and
PROD_MIX = relative proportion of parcel, registered mail, and express mail to total mail volume received by a store.

new plan. This finding suggests that it is critical to explicitly take into account the major differences in environments of business units when forming a reference group. The use of peer group performance to gain information about an individual employee's effort and to insulate employees' compensation from risks they cannot control is more valuable in contracting with a higher degree of common uncertainty.

On the other hand, the coefficients on the proxies for *UNFAIR* allow us to test the predicted relation between perceived unfairness and store performance. For all six regressions, the coefficients on *UNFAIR* are negative and significant. When $UNSATIS_{t-1}$ proxies for *UNFAIR*, the coefficient is negative and significant ($p < 0.0001$, two-tailed) as seen in columns (2), (4), and (6), providing strong evidence on our prediction that the new RPE plan's incentive effect will be mitigated in stores with likely high perceived unfairness. When *CHANGE* proxies for *UNFAIR*, the coefficient is negative and marginally significant ($p = 0.06$ to 0.07, two-tailed) as seen in columns (1), (3), and (5), providing weak evidence that the new plan's incentive effect was attenuated in stores with high perceived unfairness.

To examine the economic significance of the results, we assess the marginal effect of *COMMON* and *UNFAIR* on store profitability. For example, column (1) suggests that stores that requested to change their reference group in 2000 (*CHANGE*) exhibit lower profitability than other stores by 0.088, holding all else constant. Similarly, column (2) suggests that a one-unit increase in decile ranking of $UNSATIS_{t-1}$ is associated with a 0.038 decrease in store profitability, holding all else constant. On the other hand, a change in membership to a reference group with a higher degree of common uncertainty ($COMMON_{t-1}$) is associated with approximately a 0.01 increase in store profitability. Taken together, these results suggest that the new RPE contract's incentive effect is mitigated by perceived unfairness and that the results are robust to alternative ways of identifying stores with a high level of perceived unfairness.

TABLE 5
Regressions Examining the Association among Degree of Common Uncertainty, Perceived Unfairness, and Performance
(t = 1998–1999)

$$PROFRATE_{i,t} = \delta_0 + \delta_1 COMMON_{i,t-1} + \delta_2 UNFAIR_{i,t} + \delta_3 SERVICE\_AREA_{i,t} + \delta_4 POP_{i,t} + \delta_5 N\_BANK_{i,t} + \delta_6 METRO_{i,t} + \delta_7 DIST_{i,t} + \delta_8 N\_DELPOINT_{i,t} + \delta_9 N\_COUNTER_{i,t} + \delta_{10} N\_DELSEC_{i,t} + \delta_{11} SKILL_{i,t} + \delta_{12} PROD\_MIX_{i,t} + \varepsilon_{i,t}$$

		(1)	(2)	(3)	(4)	(5)	(6)
		Proxy for *COMMON*					
		RSQRANK		*RESIDRANK*		*DISTRANK*	
Proxy for *UNFAIR*	**Predicted Sign**	*CHANGE*	$UNSATIS_{t-1}$	*CHANGE*	$UNSATIS_{t-1}$	*CHANGE*	$UNSATIS_{t-1}$
Constant		0.252	1.837	1.117	2.974**	0.872	2.797*
		(0.15)	(1.20)	(0.67)	(1.97)	(0.52)	(1.85)
$COMMON_{t-1}$	(+)	0.008*	0.012***	0.011**	0.012**	0.004	0.010**
		(1.79)	(2.92)	(1.99)	(2.37)	(0.90)	(2.28)
UNFAIR	(−)	−0.088*	−0.038***	−0.087*	−0.037***	−0.091*	−0.038***
		(−1.84)	(−9.76)	(−1.83)	(−9.53)	(−1.90)	(−9.72)
SERVICE_AREA	?	−0.015***	−0.012***	−0.014***	−0.010**	−0.014***	−0.010**
		(−3.74)	(−3.15)	(−3.25)	(−2.56)	(−3.44)	(−2.53)
POP	?	0.104***	0.081***	0.097***	0.073***	0.094***	0.064***
		(4.46)	(3.85)	(4.24)	(3.47)	(4.03)	(3.02)
N_BANK	?	0.222***	0.176***	0.213***	0.166***	0.218***	0.171***
		(7.57)	(6.53)	(7.28)	(6.14)	(7.44)	(6.34)
METRO	?	0.190***	0.208***	0.187***	0.206***	0.192***	0.208***
		(7.84)	(9.46)	(7.68)	(9.28)	(7.88)	(9.44)
DIST	?	−0.073**	−0.034	−0.075**	−0.036	−0.069**	−0.025
		(−2.14)	(−1.10)	(−2.18)	(−1.16)	(−2.00)	(−0.80)
N_DELPOINT	?	0.001	−0.002	0.002	−0.001	0.001	−0.002
		(0.36)	(−0.88)	(0.69)	(−0.41)	(0.34)	(−1.00)
N_COUNTER	?	0.016	0.019	0.014	0.017	0.016	0.019
		(0.67)	(0.86)	(0.58)	(0.75)	(0.66)	(0.86)
N_DELSEC	?	−0.027	0.027	−0.044	0.007	−0.026	0.030
		(−0.40)	(0.44)	(−0.65)	(0.11)	(−0.38)	(0.49)

(*continued on next page*)

SKILL	?	−0.062 (−0.60)	−0.133 (−1.44)	−0.103 (−1.01)	−0.188** (−2.05)	−0.092 (−0.90)	−0.179** (−1.96)
PROD_MIX	?	1.567*** (4.19)	1.052*** (3.09)	1.621*** (4.36)	1.152*** (3.39)	1.678*** (4.46)	1.251*** (3.67)
n		410	410	410	410	410	410
Adjusted R^2		80.1%	83.9%	80.2%	83.7%	80.0%	83.7%
Change in R^2 by *COMMON* and *UNFAIR*		0.3%**	4.1%***	0.4%***	3.9%***	0.2%*	3.9%***

---

*, **, *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively (two-tailed).

Reported t-statistics are based on the White (1980) standard errors.

Variable Definitions:

RSQRANK = the adjusted R^2 ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating greater R^2;
RESIDRANK = residual dispersion ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating smaller residual dispersion;
DISTRANK = the average Euclidean distance of stores in a group from the reference group mean of performance control variables; ranges from 1 through 9 with higher values indicating smaller distance from the mean;
CHANGE = dummy variable representing stores that requested to change their reference group in 2000, taking the value 1 if the store requested a change, and 0 otherwise;
$UNSATIS_{t-1}$ = the decile ranking of the difference between a store's rank on profitability growth and on relative profitability, with higher values indicating stores that exhibited high profitability growth, but low relative profitability in year $t-1$;
SERVICE_AREA = square kilometers of the area that a store is responsible for at year t (in hundreds);
POP = the natural log of population (in thousands) of the area that a store is responsible for at year *t*;
N_BANK = number of financial institutions in the area that a store is responsible for at year *t* (in hundreds);
METRO = variable taking on the values of 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively;
DIST = average daily delivery distance per mail carrier of a store at year *t* (in hundred kilometers);
N_DELPOINT = average number of daily delivery points per mail carrier of a store at year *t* (in hundreds);
N_COUNTER = number of available counters of a store at year *t* (in hundreds);
N_DELSEC = number of delivery districts of a store at year *t* (in hundreds);
SKILL = average labor cost per employee of a store (million Korean won); and
PROD_MIX = relative proportion of parcel, registered mail, and express mail to total mail volume received by a store.

Association between Perceived Unfairness and Degree of Common Uncertainty

Panel A of Table 6 provides the results from estimating the binary logit regression that examines whether the degree of common uncertainty is negatively associated with the likelihood of predicting stores with high perceived unfairness when *CHANGE* is used as a dependent variable. Consistent with H4, when *COMMON* is measured with adjusted R^2, the coefficient on *COMMON* is negative and significant ($p = 0.002$, two-tailed). This suggests that stores are more (less) likely to suffer from employees' perceived unfairness under the new RPE incentive contract if the degree of common uncertainty is low (high) within their reference group. The result also implies that the perceived unfairness reported in 13 stores after the new contract could in part be attributable to low common uncertainty in their reference group. The coefficient on *COMMON* is also negative and significant ($p = 0.045$, two-tailed) when residual dispersion is employed as a proxy for *COMMON*. The coefficient on *COMMON* is not statistically different from zero ($p = 0.17$) when the average store-level Euclidean distance within reference group is employed as a proxy for *COMMON*.

For additional analysis, we estimate Equation (5), in which *UNSATIS* is regressed on *COMMON* using an ordered logit model. Panel B of Table 6 shows that the coefficient on *COMMON* is negative and significant ($p = 0.018$, two-tailed) when *COMMON* is measured with residual dispersion. The coefficient is also negative and significant ($p = 0.034$, two-tailed) when *COMMON* is measured with the average store-level Euclidean distance within reference group, but is not significant with adjusted R^2 as a metric for *COMMON*.

Even though the results from estimating logit regressions make it difficult to draw unambiguous inferences on the association between the degree of common uncertainty and perceived unfairness, we believe that taken together, the results of Tables 5 and 6 suggest that the net benefits of introducing the new RPE contract are conditional on the degree of common uncertainty. This is because the degree of common uncertainty appears to be inversely related to likely perceived unfairness related to introducing the contract.

Robustness Analyses

We assess the sensitivity of the primary results in a number of additional ways. Instead of tabulating the results of the additional tests, we discuss them briefly in this subsection.

First, we include the residuals obtained by annual regression of mail volume handled per employee of each store on major exogenous factors—population, service area, average delivery distance, number of financial institutions—and location of store (the *METRO* variable) as an additional explanatory variable, and re-estimate Equations (2) and (3). By doing so, we additionally control for unknown omitted factors that may influence store revenue through change in mail volume. The coefficient on *FYR_PLAN* and *SYR_PLAN* remains highly significant (t-statistic = 7.49 and 13.72, respectively, two-tailed) and all the proxies of common uncertainty and perceived unfairness continue to exhibit expected signs except that the coefficient on *RESIDRANK* becomes insignificant.

Second, instead of *PROFRATE*, we use operating profit scaled by square feet of a store as an alternative dependent variable and repeat our analysis on the impact of change in incentive plan. The coefficient on *FYR_PLAN* and *SYR_PLAN* remains highly significant (t-statistic = 5.10 and 7.92, respectively, two-tailed).

Third, following prior literature (Ittner and Larcker 1998b; Banker et al. 2000), we include prior performance ($PROFRATE_{t-1}$) as a control instead of including various performance controls described in Section V. The coefficient on *FYR_PLAN* and *SYR_PLAN* remains highly significant (t-statistic = 8.60 and 7.26, respectively, two-tailed) and our inferences remain unchanged.

TABLE 6
Logit Regressions Examining the Association between Degree of Common Uncertainty and Perceived Unfairness
(t = 1998–1999)

Panel A: Logit Regressions Using *CHANGE* as Dependent Variable

$$Prob(CHANGE = 1) = F\Big(\alpha_0 + \alpha_1 COMMON_{i,t} + \alpha_2 SERVICE\_AREA_{i,t} + \alpha_3 POP_{i,t} + \alpha_4 N\_BANK_{i,t} + \alpha_5 METRO_{i,t} + \alpha_6 N\_EMP_{i,t}$$
$$+ \alpha_7 DIST_{i,t} + \alpha_8 N\_DELPOINT_{i,t} + \alpha_9 N\_COUNTER_{i,t} + \alpha_{10} N\_DELSEC_{i,t} + \sum_{i=1}^{7} \alpha_{i+10} REGION_i\Big)$$

		(1)		(2)		(3)	
		Proxy for *COMMON*					
		RSQRANK		*RESIDRANK*		*DISTRANK*	
Variable	Predicted Sign	Coefficient	Wald Chi-Square	Coefficient	Wald Chi-Square	Coefficient	Wald Chi-Square
Constant		1.073	0.74	0.278	0.05	−2.150**	4.19
COMMON	(−)	−0.324***	9.69	−0.232**	4.03	0.113	1.87
SERVICE_AREA	?	−0.002**	4.54	−0.001	2.03	−0.001	1.25
POP	?	−0.009***	7.52	−0.007**	5.18	−0.006*	3.50
N_BANK	?	−0.012	1.54	−0.012	1.50	−0.011	1.26
METRO	?	−0.179	0.11	−0.201	0.15	−0.063	0.01
N_EMP	?	0.024*	2.93	0.024*	3.25	0.027**	3.87
DIST	?	−0.006	0.65	−0.006	0.62	−0.006	0.52
N_DELPOINT	?	0.001	2.60	0.001*	3.44	0.0005	2.14
N_COUNTER	?	−0.113*	3.44	−0.100	2.68	−0.100	2.64
N_DELSEC	?	0.014	0.35	0.004	0.03	−0.001	0.002
n		410		410		410	
Pseudo R^2		15.0%		11.3%		10.1%	
% Correctly Predicted		81.0%		74.6%		75.6%	
Likelihood ratio χ^2		27.37		20.71		18.42	

(*continued on next page*)

TABLE 6 (Continued)

Panel B: Ordered Logit Regressions Using *UNSATIS* as Dependent Variable

$$Prob(UNSATIS) = F\Bigg(\delta_0 + \delta_1 COMMON_{i,t} + \delta_2 SERVICE\_AREA_{i,t} + \delta_3 POP_{i,t} + \delta_4 N\_BANK_{i,t} + \delta_5 METRO_{i,t} + \delta_6 N\_EMP_{i,t}$$

$$+ \delta_7 DIST_{i,t} + \delta_8 N\_DELPOINT_{i,t} + \delta_9 N\_COUNTER_{i,t} + \delta_{10} N\_DELSEC_{i,t} + \sum_{i=1}^{7} \delta_{i+10} REGION_i\Bigg)$$

		(1)		(2)		(3)	
				Proxy for *COMMON*			
		RSQRANK		*RESIDRANK*		*DISTRANK*	
Variable	**Predicted Sign**	**Coefficient**	**Wald Chi-Square**	**Coefficient**	**Wald Chi-Square**	**Coefficient**	**Wald Chi-Square**
COMMON	(−)	0.041	1.01	−0.113**	5.53	−0.084**	4.45
SERVICE_AREA	?	0.001***	9.58	0.001***	10.43	0.001***	11.91
POP	?	−0.001	1.16	−0.002**	4.20	−0.002*	2.91
N_BANK	?	−0.012***	16.32	−0.012***	15.68	−0.012***	16.63
METRO	?	0.646***	11.22	0.579***	8.70	0.592***	9.03
N_EMP	?	−0.007	2.37	−0.010**	4.62	−0.008*	3.21
DIST	?	0.007**	5.64	0.008***	7.34	0.008***	7.62
N_DELPOINT	?	−0.0003*	3.54	−0.0003*	3.58	−0.0004**	4.38
N_COUNTER	?	0.0005	0.06	0.001	0.10	0.001	0.09
N_DELSEC	?	0.016*	3.22	0.020**	5.44	0.019**	4.82
n		410		410		410	
Generalized R^2		37.4%		38.1%		37.9%	
% Correctly Predicted		72.6%		72.9%		73.0%	
Likelihood ratio χ^2		192.11		196.60		195.55	

(continued on next page)

---

*, **, and *** Significant at the 10 percent, 5 percent, and 1 percent levels, respectively (two-tailed).

For parsimony, the nine intercepts produced by the ordered logit model and the coefficients on region-dummies are not reported.

CHANGE = dummy variable representing stores that requested to change their reference group in 2000, taking the value 1 if the store requested a change, and 0 otherwise;

UNSATIS = the decile ranking of the difference between a store's rank on profitability growth and on relative profitability, with higher values indicating stores that exhibited high profitability growth, but low relative profitability;

RSQRANK = the adjusted R^2 ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating greater R^2;

RESIDRANK = residual dispersion ranking of a store obtained by running reference group-specific regression Equation (1); ranges from 1 through 9 with higher values indicating smaller residual dispersion;

DISTRANK = the average Euclidean distance of stores in a group from the reference group mean of performance control variables; ranges from 1 through 9 with higher values indicating smaller distance from the mean;

SERVICE_AREA = square kilometers of the area that a store is responsible for at year *t* (in hundreds);

POP = the natural log of population (in thousands) of the area that a store is responsible for at year *t*;

N_BANK = number of financial institutions in the area that a store is responsible for at year *t* (in hundreds);

METRO = variable taking on the values of 1, 2, 3, and 4 for rural area, small-sized city, middle-sized city, and metropolitan area, respectively;

DIST = average daily delivery distance per mail carrier of a store at year *t* (in hundred kilometers);

N_DELPOINT = average number of daily delivery points per mail carrier of a store at year *t* (in hundreds);

N_COUNTER = number of available counters of a store at year *t* (in hundreds); and

N_DELSEC = number of delivery districts of a store at year *t* (in hundreds);

Fourth, we re-estimate Equation (3) using year 1999 data only instead of 1998 and 1999 data in order to allow for the possibility that it may take considerable time before the perceptions of common uncertainty and unfairness form and influence performance. Our inferences regarding H2 and H3 remain qualitatively unchanged except that the coefficients on *CHANGE* become insignificant.

Fifth, we estimate Equations (2) and (3) using standard ordinary least square (OLS) after including the seven *REGION* dummy variables to control for unspecified geographical region-specific differences in stores' profitability. The estimation results are very similar to those obtained by fixed-effects regressions. We also estimate Equations (3), (4), and (5) using the raw values of our proxies for *COMMON* in place of rank values and our inferences remain qualitatively unchanged, suggesting that using the rank value does not confound our findings.

Sixth, we estimate Equation (3) after we include a year 1999 dummy variable to control for a potential "learning effect" in the second year following the new plan. Our inferences are not changed.

Finally, we estimate Equations (3) and (4) using the following two alternative ways of measuring the level of perceived unfairness (*UNFAIR*). As a variation of *UNSATIS*, we employ a dummy variable representing stores that exhibited top 50 percent profitability growth, but bottom 50 percent relative profitability. As a variation of *CHANGE*, we use a dummy variable for the 37 stores that in November 1999 had their reference group changed to a "lower" reference group relative to their first reference group in December 1997. Our inferences remain unchanged.

VII. CONCLUDING REMARKS

Using annual performance evaluation data of 214 stores of a postal service provider in Korea during 1997 through 1999, we find that adding relative performance measures to an incentive plan is positively associated with financial performance at the store level and that under the new incentive plan, the degree of common uncertainty is positively associated with store profitability. Furthermore, we find evidence that the new plan's incentive effect is mitigated in stores at which the employees' perceived unfairness is likely to be high. Finally, we find that the net benefits of introducing the new contract may be conditional on the degree of common uncertainty because the degree of common uncertainty appears to be inversely related to perceived unfairness related to the new contract, with its introduction of relative performance measures.

Our results are especially striking given that the postal service provider did not switch from a pure non-RPE incentive plan to an RPE plan. Rather, both plans had an element of RPE through a tournament structure in which only employees in stores that scored in the top 50 percent on the designated performance measure received a bonus. Under the new incentive plan, the postal service classified all stores into nine reference groups, based on business environments. The new plan also introduced a performance measure that placed comparatively large weights on profitability (i.e., store revenue divided by store operating cost) relative to average reference-group profitability, and on productivity (i.e., mail volume per store employee) relative to average reference-group productivity.

Our results have several implications for implementation of RPE incentive plans. First, even though an RPE incentive plan appears to induce agents to exert more effort by reducing risk and introducing competition, a significant level of dysfunctional responses can occur, for example, due to decreased morale or skepticism brought on by employees' perceived unfairness of benchmarked targets. Second, the observed perceived unfairness caused by introducing an RPE contract may mitigate the incentive effect. Third, our study suggests that the level of perceived unfairness increases as the degree of common uncertainty decreases. This suggests that careful attention should be directed to setting the benchmark standards against which the performance of business units is compared, and it is critical to explicitly take into account the major differences in environments of business units when forming a reference group. This supports conventional wisdom that benchmark units should be from similar environments (Dopuch and Gupta 1997; Brickley et al. 2003).

This study's findings are subject to several important caveats. First, this study uses observations from a single firm and it is an open question whether we can generalize results from a single research site. Second, although we perform a number of robustness tests, the absence of a natural control sample and long time-series data make it difficult to completely rule out all alternative explanations for our findings. Third, we are unable to measure *ex ante* perceived unfairness, i.e., employees' perception of the incentive plan as unfair *before* the bonuses were granted. Instead, we develop *ex post* proxies that attempt to measure perceived unfairness *after* the bonuses were granted, but our measure may in part also capture dissatisfaction with not receiving a bonus. Future research could attempt to replicate our findings in a setting that controls for our limitations.

REFERENCES

Aggarwal, R., and A. Samwick. 1999. Executive compensation, strategic competition, and relative performance evaluation: Theory and evidence. *Journal of Finance* 54: 1999–2043.

Albuquerque, A. 2005. Who are your peers? A study of relative performance evaluation. Working paper, University of Rochester.

Anderson, M., R. Banker, and S. Ravindran. 2001. Another look at relative performance evaluation hypothesis. Working paper, The University of Texas at Dallas.

Antle, R., and A. Smith. 1986. An empirical investigation of relative performance evaluation of corporate executives. *Journal of Accounting Research* 24: 1–39.

Baek, J., J. Kang, and K. Park. 2004. Corporate governance and firm value: Evidence from the Korean financial crisis. *Journal of Financial Economics* 71: 265–313.

Banker, R., S. Lee, and G. Potter. 1996. A field study of the impact of a performance-based incentive plan. *Journal of Accounting and Economics* 21: 195–226.

———, G. Potter, and D. Srinivasan. 2000. An empirical investigation of an incentive plan that includes nonfinancial measures. *The Accounting Review* 75: 65–92.

Belsley, D., E. Kuh, and R. Welsch. 1980. *Regression Diagnostics*. New York, NY: John Wiley & Sons.

Brickley, J., C. Smith, and J. Zimmerman. 2003. *Managerial Economics and Organizational Architecture*. 3rd edition. Chicago, IL: R. D. Irwin, Inc.

Brockner, J., and B. Wiesenfeld. 1996. An integrative framework for explaining reactions to decisions: Interactive effects of outcomes and procedures. *Psychological Bulletin* 120: 189–208.

Casas-Arce, P., and F. Martínez-Jerez. 2005. Relative performance compensation, contexts and dynamic incentives. Working paper, Harvard University.

Cropanzano, R., and R. Folger. 1989. Referent cognitions and task decision autonomy: Beyond equity theory. *Journal of Applied Psychology* 74: 293–299.

Crotty, J., and K. Lee. 2001. Economic performance in post-crisis Korea: A critical perspective on neoliberal restructuring. Working paper, University of Massachusetts Amherst.

Dopuch, N., and M. Gupta. 1997. Estimation of benchmark performance standards: An application of public school expenditures. *Journal of Accounting and Economics* 23: 141–161.

Fisher, J., S. Peffer, and G. Sprinkle. 2003. Budget-based contracts, budget levels, and group performance. *Journal of Management Accounting Research* 15: 51–74.

Frederickson, J. 1992. Relative performance information: The effects of common uncertainty and contract type on agent effort. *The Accounting Review* 67: 647–669.

Frees, E. 2004. *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. New York, NY: Cambridge University Press.

Gibbons, R., and K. Murphy. 1990. Relative performance evaluation for chief executive officers. *Industrial and Labor Relation Review* 43: 30–51.

Gu, Z. 2002. Cross-sample incomparability of R²s and additional evidence on value-relevance changes over time. Working paper, Carnegie Mellon University.

Hansen, S., D. Otley, and W. Van der Stede. 2003. Practice development in budgeting: An overview and research perspective. *Journal of Management Accounting Research* 15: 95–116.

Holmstrom, B. 1982. Moral hazard in teams. *Bell Journal of Economics* 13: 324–340.

Indjejikian, R., and D. Nanda. 1999. Dynamic incentives and responsibility accounting. *Journal of Accounting and Economics* 27: 177–201.

Ittner, C., and D. Larcker. 1998a. Innovations in performance measurement: Trends and research implications. *Journal of Management Accounting Research* 10: 205–238.

———, and ———. 1998b. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research* 36 (Supplement): 1–35.

Janakiraman, S., R. Lambert, and D. Larcker. 1992. An empirical investigation of the relative performance evaluation hypothesis. *Journal of Accounting Research* 30: 53–69.

Johnson, R., and D. Wichern. 2002. *Applied Multivariate Statistical Analysis*. 5th edition. Upper Saddle River, NJ: Prentice Hall.

Kennedy, P. 1992. *A Guide to Econometrics*. Cambridge, MA: MIT Press.

Kren, L. 2002. Common uncertainty effects on the use of relative performance evaluation for corporate chief executives. *Advances in Accounting* 19: 119–138.

Lazear, E. 2000. Performance pay and productivity. *American Economic Review* 90: 1346–1361.

Libby, T. 2001. Referent cognitions and budgetary fairness: A research note. *Journal of Management Accounting Research* 13: 91–105.

Magner, N., R. Welker, and T. Campbell. 1995. The interactive effect of budgetary participation and budget favorability on attitudes towards budgetary decision makers: A research note. *Accounting, Organizations and Society* 20: 611–618.

Maruyama, S., and T. Nakajima. 2002. *Efficiency Measurement and Productivity Analysis for Japanese Postal Service*. Tokyo, Japan: Institute for Posts and Telecommunications Policy.

———, and ———. 2003. The productivity analysis of postal services: Global comparison of the technical efficiency and the total factor productivity. In *Postal and Delivery Services: Delivering on Competition*, edited by M. Crew, and P. Kleindorfer. Berlin, Germany: Springer.

McFarlin, D., and P. Sweeney. 1992. Distributive and procedural justice as predictors of satisfaction with personal and organizational outcomes. *Academy of Management Journal* 35: 626–638.

Merchant, K., and J. Manzoni. 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review* 64: 539–558.

Milgrom, P., and J. Roberts. 1992. *Economics, Organization and Management*. Englewood Cliffs, NJ: Prentice Hall.

Mookherjee, D. 1984. Optimal incentive schemes with many agents. *Review of Economic Studies* 51: 433–446.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* 37: 7–63.

Seta, J. 1982. The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology* 42: 281–291.

Sloan, R. 1993. Accounting earnings and top executive compensation. *Journal of Accounting and Economics* 16: 55–100.

Sprinkle, G. 2000. The effect of incentive contracts on learning and performance. *The Accounting Review* 75: 299–326.

White, H. 1980. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica* 48: 817–838.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 567–588

Does Recognition versus Disclosure Matter? Evidence from Value-Relevance of Banks' Recognized and Disclosed Derivative Financial Instruments

Anwer S. Ahmed
Texas A&M University

Emre Kilic
University of Houston

Gerald J. Lobo
University of Houston

ABSTRACT: We provide evidence on how investor valuation of derivative financial instruments differs depending upon whether the fair value of these instruments is *recognized* or *disclosed*. Expanded disclosures and accounting practices prior to SFAS No. 133 and mandatory recognition of derivative fair values after SFAS No. 133 provide a natural setting for comparing the valuation implications of recognized and disclosed derivative fair value information. This unique setting mitigates many of the research design problems with recognition versus disclosure studies. Using a sample of banks that simultaneously hold recognized and disclosed derivatives prior to SFAS No. 133, we find that the valuation coefficients on recognized derivatives are significant, whereas the valuation coefficients on disclosed derivatives are not significant. Further, using a sample of banks that have only disclosed derivatives prior to SFAS No. 133, which are recognized after SFAS No.133, we find that while the valuation coefficients on disclosed derivatives are not significant, the valuation coefficients on recognized derivatives are significant. These results are consistent with the view that recognition and disclosure are not substitutes. Our findings suggest that SFAS No. 133 has increased the transparency of derivative financial instruments.

Keywords: *SFAS No. 133; recognition; disclosure; derivative financial instruments.*

Data Availability: *All data are available from public sources.*

We thank Rashad Abdel-khalik, Dan Dhaliwal (the editor), Steve Fortin, David Harris, Mark Kohlbeck, Thomas Lys, Michel Magnan, Mary Lea McAnally, Katherine Schipper, Terry Shevlin, Terry Warfield, Ross Watts, two anonymous referees, workshop participants at Texas A&M University, University of Wisconsin–Madison, Concordia University, Syracuse University, the 2004 AAA Annual Meeting and the 2005 FARS Midyear Meeting for helpful comments.

Editor's note: This paper was accepted by Dan Dhaliwal.

Submitted October 2004
Accepted November 2005

JEL Classification: *G14; G21; M41.*

I. INTRODUCTION

We provide evidence on how investor valuation of derivative financial instruments differs depending upon whether the fair value of these instruments is *recognized* or *disclosed*.[1] We find that in a period prior to SFAS No. 133 (FASB 1998), the valuation coefficients on recognized derivatives are statistically significant, whereas the valuation coefficients on disclosed derivatives are not significant. Furthermore, for a sample of banks that held only disclosed derivatives before SFAS No. 133, we find that the valuation coefficients on disclosed derivatives are not significant before SFAS No. 133 (when they were disclosed), but are significant after SFAS No. 133 (when they were recognized).

The research question we address is important for a number of reasons. First, from a standard-setting perspective, recognition and disclosure are viewed as alternative accounting treatments. Therefore, whether recognition versus disclosure affects users' decisions has been an issue of considerable interest to standard setters, practicing professionals, and academic researchers. In efficient markets with minimal information-processing costs, one would not expect the accounting treatment (i.e., recognition versus disclosure) to matter. However, if there are (1) costs of processing information (Barth et al. 2003), (2) systematic biases in how investors process information, such as limited attention (Hirshleifer and Teoh 2003), or (3) differences in perceived reliability of recognized versus disclosed items, this choice can matter to users.

Second, although accounting for derivatives has been one of the most widely debated issues among academics and practitioners, to our knowledge, no prior study examines explicitly whether investor valuation of derivatives depends upon recognition versus disclosure in the pre-SFAS No. 133 period and no study has examined whether SFAS No. 133 has affected investor valuation of derivatives.

We perform two complementary tests of differences in investor valuation of *recognized* versus *disclosed* derivative instruments on two samples of bank holding companies. The first test examines whether recognized derivative instruments are valued differently from disclosed derivative instruments for a sample of banks that simultaneously hold both recognized and disclosed derivative financial instruments for risk management purposes in a pre-SFAS No. 133 period. We find that valuation coefficients are significantly greater than zero for recognized derivative instruments but not for disclosed derivative instruments.

The second test examines whether there is a difference between valuation of *recognized* derivatives in the post-SFAS No.133 period and valuation of *disclosed* derivatives in the pre-SFAS No. 133 period. We find that the valuation coefficients on the recognized derivatives (in the post-SFAS No.133 period) are significantly positive, whereas the valuation coefficients on the disclosed derivatives (in the pre-SFAS No. 133 period) are not significantly different from zero.

We contribute to two streams of prior research. First, a number of prior studies investigate implications of recognition versus disclosure of other accounting items (e.g., stock options and post-retirement benefits). Typically, these studies use a setting where some firms disclose while others recognize (e.g., Aboody 1996), or where a new standard requires recognition of a previously disclosed item (e.g., Davis-Friday et al. 1999). The studies in the first group are subject to self-selection bias because firms self-select into recognizers or disclosers (Bernard and Schipper 1994). One potential problem with the studies in the second group is that the amount and/or nature of the information reported about the item is rarely held constant because the same firm could be disclosing and recognizing at different points in time (Bernard and Schipper 1994).

---

[1] Banks hold derivative financial instruments to meet customer demands, to trade these products for their own accounts (for trading purposes), and to manage risk (for nontrading purposes). The focus of this paper is on derivative financial instruments held for risk management purposes. We use "risk management derivative financial instruments" and "derivative financial instruments" interchangeably in the rest of the paper. All trading derivatives are already recognized at fair value with any realized or unrealized gains or losses included in current earnings. See the Appendix for examples of how banks in our sample define their trading and nontrading derivative activities in footnotes to their financial statements.

Our pre-SFAS No. 133 tests employ a within-firm design that mitigates these typical problems in recognition versus disclosure studies because banks could report both recognized and disclosed risk management derivatives at the *same* point in time. Our pre-SFAS 133 versus post-SFAS No. 133 analysis, like other studies that examine effects of changes in accounting standards, is potentially subject to the second problem. However, this problem is likely to be considerably less serious in our setting because SFAS No. 133 explicitly requires derivative fair values to be measured using the *same* valuation guidance and techniques as those used in the pre-SFAS No. 133 period. In other words, the amount and/or nature of information in the post-SFAS No. 133 period is likely to be very similar to the amount and/or nature of information in the pre-SFAS No. 133 period.

In addition to these general design advantages offered by the setting we use, our paper specifically differs from other recognition versus disclosure studies in another important respect. Those studies employ a "levels" analysis only, whereas we provide consistent evidence using both "changes" and "levels" analyses. Because the "changes" specification is less subject to omitted variables bias, our evidence is more robust and allows stronger inferences to be made about recognition versus disclosure. Our study also differs from prior recognition versus disclosure studies by examining the same research question using two different samples with two different research designs, i.e., pre-SFAS No. 133 analysis, and pre- versus post-SFAS No. 133 analysis. Second, our study is related to the line of research that examines the value-relevance of disclosed information about financial instruments. While prior studies on fair value disclosures of derivatives generally find that the disclosures are not value-relevant (e.g., Barth et al. 1996; Eccher et al. 1996; Nelson 1996), using disclosures required by SFAS. No. 119 (FASB 1994), Venkatachalam (1996) finds that banks' derivative fair value disclosures are value-relevant. However, because his study is designed to test the valuation implications of derivative fair values rather than recognition versus disclosure of these amounts, he aggregates fair values of disclosed and recognized derivative financial instruments and restricts their coefficients to be the same. We allow the coefficient estimates on recognized and disclosed risk management derivatives to differ and document significant differences. Furthermore, we provide evidence on whether SFAS No. 133 has had an impact on the value-relevance of derivatives.

While our findings are consistent with recent limited attention or costly information-processing models, an alternative explanation for our findings is that there are differences in perceived reliability of recognized and disclosed derivatives due to potentially higher standards imposed by audit firms and management for recognized items than for disclosed items.[2]

The rest of this paper is organized as follows: Section II describes the institutional background and financial reporting environment related to derivative financial instruments before and after the adoption of SFAS No. 133. Section III presents the research design and empirical methods employed in our study. Section IV presents the results of our empirical analysis. Section V provides a summary and conclusion.

---

[2] For example, Libby et al. (2005) find in an experimental setting that "audit partners require greater correction of misstatements in recognized amounts than in equivalent disclosed amounts."

II. INSTITUTIONAL BACKGROUND

Accounting for Derivative Financial Instruments Prior to SFAS No. 133

Prior to SFAS No. 133, accounting for derivatives was governed by (1) SFAS Nos. 52 and 80 (FASB 1981, 1984), and (2) consensus positions of the Emerging Issues Task Force (EITF). Additional accounting guidance for banks was provided by the *Audit and Accounting Guide for Banks and Savings Institutions* (AICPA 2000).[3] Accounting for derivative instruments that were not specifically covered by SFAS Nos. 52 and 80 developed largely by analogy to these standards. The accounting treatment for derivative instruments under the requirements of SFAS No. 52 and 80 could be categorized into: (1) hedge accounting, (2) synthetic alteration (or settlement accounting), and (3) speculative accounting.

If a derivative instrument was being used to hedge an existing asset, liability, or anticipated transaction, then it was accounted for in a manner consistent with the requirements of hedge accounting. On the other hand, if a derivative instrument was being used to change the nature of one financial instrument to that of another financial instrument (such as an interest-rate swap changing floating-rate debt to fixed-rate debt), it was accounted for in accordance with the requirements of synthetic alteration (or settlement accounting).

Under both hedge and settlement accounting, the accounting treatment for a derivative instrument was determined by the accounting treatment for the underlying asset or liability to which the derivative instrument is designated. Specifically, a derivative instrument was recognized at fair value on the balance sheet as an asset or a liability if the related underlying asset or liability was also carried at fair value or lower of cost or fair value (such as a foreign-exchange forward hedging the foreign currency risk of an available-for-sale security).[4] On the other hand, if a derivative instrument was hedging an asset or liability carried at historical or amortized cost, it was not recognized at fair value on the balance sheet (such as an interest rate swap hedging a variable-rate debt). Depending on the underlying asset, liability, or transaction, either no entry was made or the derivative instrument was carried at historical cost, which often equaled zero.

If a derivative instrument was used for trading or speculative purposes, it was accounted for under speculative accounting, which required the derivative instrument to be recognized at fair value.

To summarize, depending on the accounting for the underlying asset, liability, or transaction, certain derivatives were reported in financial statements at fair value under the accounting rules for derivatives and hedging activities prior to SFAS No. 133, and others were not.[5] This created a financial reporting environment in which it was possible for a bank to report some of its risk management derivatives on the balance sheet at fair value, and to just disclose the rest of its risk management derivatives, i.e., hold recognized and disclosed derivative instruments simultaneously.

---

[3] Specifically, EITF Issue 84-07 on *Termination of Interest Rate Swaps*, EITF Issue 84-36 on *Interest Rate Swap Transactions*, EITF Issue 88-08 on *Mortgage Swaps*, EITF Issue 90-17 on *Hedging Foreign Risk with Purchased Options Contracts*, and EITF Issue 91-04 *on Hedging Foreign Currency Risks with Complex Options* were the EITF issues that provided guidance for accounting for derivative financial instruments (EITF 1984a, 1984b, 1988, 1990, 1991).

[4] Regarding these types of derivative instruments, SFAS No. 80 (in paragraph 5) states: "If an enterprise includes unrealized changes in the fair value of a hedged item in income, a change in the market value of the related futures contract shall be recognized in income when the change occurs. The same accounting shall be applied to a futures contract that hedges an anticipated transaction if the asset to be acquired or liability to be incurred will be reported at fair value subsequent to acquisition or incurrence."

Requirements under SFAS No. 107 (FASB 1991) made fair values of all derivative financial instruments available to investors, while SFAS No. 119 required banks to clearly indicate whether the aggregate fair value of the derivatives portfolio represents a net asset or a net liability position and to provide disaggregated information on carrying amounts, fair values, and contractual amounts of derivatives classified by category of instrument and by purpose for which they are held.[6] These expanded disclosures and accounting practices prior to SFAS No. 133 provide a natural setting for comparing the valuation implications of recognized and disclosed derivative fair value information.

Accounting for Derivative Financial Instruments after SFAS No. 133

Statement of Financial Accounting Standards No. 133, *Accounting for Derivative Instruments and Hedging Activities* was issued in June 1998 and was effective for fiscal years beginning after June 15, 2000.[7] SFAS No. 133 requires all derivatives, without exception and regardless of the accounting treatment for the underlying asset, liability, or transaction, to be recognized in the balance sheet as either liabilities or assets. However, accounting for the gain or loss on a derivative instrument depends on how it is used. Depending upon whether it is a fair value or cash flow hedge, the gain or loss is recognized either in current earnings or comprehensive income. While SFAS Nos. 105 (FASB 1990), 107, and 119 aim at improving risk-related disclosures for financial instruments, including derivatives, SFAS No. 133 constitutes the recognition stage.

III. RESEARCH DESIGN

Overview

We use both levels and changes specifications to test the value-relevance of recognized and disclosed fair values of derivative instruments. The models assume that the relation between derivative value and firm value is cross-sectionally uniform. Furthermore, there may be omitted variables if there are unrecognized assets or liabilities, e.g., intangibles, which are priced by investors. We use the changes specification to help mitigate potential omitted variables bias (Kothari and Zimmerman 1995).

Pre-SFAS No. 133 Analysis—Levels Model

Our cross-sectional equity framework follows prior research (e.g., Ahmed and Takeda 1995; Barth et al. 1996; Venkatachalam 1996). The levels specification expresses firm value as a function of the fair values of recognized and disclosed derivatives, on-balance-sheet non-derivative assets and liabilities, and proxies for intangibles.

---

[5] SFAS No. 119 and SFAS No. 133, in addition to SFAS No. 80, acknowledge differential reporting treatment for risk management derivatives as well. SFAS No. 119 (in paragraph 48) notes: "The Board believes that some derivative financial instruments may be measured at fair value but not considered held or issued for *trading purposes*. For example, some hedging activities may entail the use of foreign exchange contracts that fail to meet the criteria for deferral accounting under FASB Statement No. 52, *Foreign Currency Translation*. Those derivative financial instruments, although measured at fair value, may be categorized as held or issued for *purposes other than trading*" (emphasis in original). SFAS No. 133 (in paragraph 234) also notes: "Under the varied accounting practices that existed before the issuance of this Statement, some derivatives were recognized in financial statements, others were not."

[6] The initial year of application for SFAS No. 107 was 1991, and the initial year of application for SFAS No. 119 was 1994.

[7] SFAS No. 133 was originally effective for the fiscal years beginning after June 15, 1999. As a result of receiving several complaints from business entities, the FASB delayed the effective date to fiscal years beginning after June 15, 2000.

In order to mitigate the effects of potential omitted variables, we add two control variables to the model. First, consistent with Barth et al. (1996), we include nonperforming loans (*NPL*) as a control variable that contains information on default risk incremental to allowance for loan loss reserves. Second, based on prior banking research (Flannery 1982; Flannery and James 1984), we include *CORE*, which is defined as domestic deposits minus time deposits, as a proxy for core deposit intangible assets. Inclusion of *CORE* as a control variable is based on the assumption that long-term relationships with depositors add significant value to banks and this value is not reflected in reported financial statement numbers.

We permit the intercept in Equation (1) to vary by year and bank in order to account for period-specific economic effects and firm-specific effects that are not captured by the explanatory variables. We use per share data in our cross-sectional levels models and check the sensitivity of our results to the use of book value of equity as a deflator.

The regression model to test whether there are differences in investor valuation of recognized and disclosed derivatives in the pre-SFAS No. 133 period is as follows:

$$\begin{aligned}\frac{MVE_{it}}{CSO_{it}} = {} & \alpha_0 + \sum_{t=1}^{T-1} \alpha_{01,t} Y_t + \sum_{i=1}^{K-1} \alpha_{02,i} BHC_i + \alpha_1 \frac{FVHA_{it}}{CSO_{it}} + \alpha_2 \frac{FVFL_{it}}{CSO_{it}} \\ & + \alpha_3 \frac{NFA_{it}}{CSO_{it}} + \alpha_4 \frac{NFL_{it}}{CSO_{it}} + \alpha_5 \frac{RNDP_{it}}{CSO_{it}} + \alpha_6 \frac{DNDP_{it}}{CSO_{it}} + \alpha_7 \frac{NPL_{it}}{CSO_{it}} \qquad (1) \\ & + \alpha_8 \frac{CORE_{it}}{CSO_{it}} + \varepsilon_{it}\end{aligned}$$

where T is the number of years in the sample period; Y_t equals 1 if the observation belongs to year t, and 0 otherwise; K is the number of bank holding companies in the sample; and BHC_i equals 1 if the observation belongs to bank holding company i, and 0 otherwise. For firm i at time t, MVE_{it} is market value of equity at the end of year t and CSO_{it} is the number of common shares outstanding. $FVFA_{it}$ and $FVFL_{it}$ are the aggregated fair values of financial assets and financial liabilities other than fair values of risk management derivative financial instruments, respectively. $FVFA_{it}$ is computed as the sum of the fair values of cash and equivalents, investment securities, net loans, and other financial assets excluding risk management derivative instruments, and $FVFL_{it}$ is computed as the sum of the fair values of deposits, short-term and long-term debt, and other financial liabilities excluding risk management derivative instruments. NFA_{it} and NFL_{it} are the aggregated book values of nonfinancial assets and nonfinancial liabilities, respectively.[8] *RNDP* and *DNDP* are the net fair values of recognized and disclosed derivatives, respectively. NPL_{it} is nonperforming loans, and $CORE_{it}$ is domestic deposits minus time deposits.

Coefficients of interest in Equation (1) are α_5 and α_6. Significant positive values for α_5 and α_6 indicate the value-relevance of fair values of recognized and disclosed derivative instruments, respectively.

Pre-SFAS No. 133 Analysis—Changes Model

In order to mitigate potential omitted variables bias, we use the following changes specification:

---

[8] We use book values of nonfinancial assets and nonfinancial liabilities because their fair values are not observable.

$$\begin{aligned}\frac{\Delta MVE_{it}}{MVE_{it-1}} &= \beta_0 + \sum_{t=1}^{T-1} \beta_{01,t}Y_t + \sum_{i=1}^{K-1} \beta_{02,t}BHC_t + \beta_1 \frac{\Delta FVFA_{it}}{MVE_{it-1}} + \beta_2 \frac{\Delta FVFL_{it}}{MVE_{it-1}} \\ &\quad + \beta_3 \frac{\Delta NFA_{it}}{MVE_{it-1}} + \beta_4 \frac{\Delta NFL_{it}}{MVE_{it-1}} + \beta_5 \frac{\Delta RNDP_{it}}{MVE_{it-1}} + \beta_6 \frac{\Delta DNDP_{it}}{MVE_{it-1}} \\ &\quad + \beta_7 \frac{\Delta NPL_{it}}{MVE_{it-1}} + \beta_8 \frac{\Delta CORE_{it}}{MVE_{it-1}} + \varepsilon_{it}\end{aligned} \tag{2}$$

We expect β_5 and β_6 in Equation (2) to have significant positive values if recognized and disclosed fair values of derivatives are value-relevant to investors.

Pre- versus Post-SFAS No. 133 Analysis

In SFAS No. 133, the FASB states (in paragraph 238):

> [SFAS No. 133] increases the visibility, comparability, and understandability of the risks associated with derivatives by requiring that all derivatives be reported as assets or liabilities and measured at fair value.

Our pre- versus post-SFAS No. 133 analysis is designed to test whether mandatory recognition of derivative instruments at fair value, as required by SFAS No. 133, has had the desired impact.

We use both levels and changes models to examine the value-relevance of derivative financial instruments, whose fair values are disclosed prior to the adoption of SFAS No. 133 and recognized subsequent to the adoption of SFAS No. 133. While these models are generally similar to the models in Equations (1) and (2), they differ from those models in two important ways. First, the coefficients on the explanatory variables are permitted to differ across the pre- and post-SFAS No. 133 periods. Second, instead of including both recognized and disclosed derivative fair values as independent variables, these models include only disclosed derivative fair values in the pre-SFAS No. 133 period and only recognized derivative fair values in the post-SFAS No. 133 period. In both periods, derivative fair values are estimated following SFAS No. 107.[9]

The levels and changes models are specified as follows:

$$\begin{aligned}\frac{MVE_{it}}{CSO_{it}} &= \lambda_0 + \sum_{t=1}^{T-1} \lambda_{01,t}Y_t + \sum_{i=1}^{K-1} \lambda_{02,i}BHC_i + \lambda_{1,1} \frac{FVFA_{it}}{CSO_{it}} \\ &\quad + \lambda_{1,2}POST * \frac{FVFA_{it}}{CSO_{it}} + \lambda_{2,1} \frac{FVFL_{it}}{CSO_{it}} + \lambda_{2,2}POST * \frac{FVFL_{it}}{CSO_{it}} \\ &\quad + \lambda_{3,1} \frac{NFA_{it}}{CSO_{it}} + \lambda_{3,2}POST * \frac{NFA_{it}}{CSO_{it}} + \lambda_{4,1} \frac{NFL_{it}}{CSO_{it}} \\ &\quad + \lambda_{4,2}POST * \frac{NFL_{it}}{CSO_{it}} + \lambda_5 \frac{DER_{it}}{CSO_{it}} + \lambda_6 POST * \frac{DER_{it}}{CSO_{it}} + \lambda_{7,1} \frac{NPL_{it}}{CSO_{it}} \\ &\quad + \lambda_{7,2}POST * \frac{NPL_{it}}{CSO_{it}} + \lambda_{8,1} \frac{CORE_{it}}{CSO_{it}} + \lambda_{8,2}POST * \frac{CORE_{it}}{CSO_{it}} + \varepsilon_{it}\end{aligned} \tag{3}$$

---

[9] SFAS No. 133 (in paragraph 243) states: "This Statement relies on the valuation guidance provided in Statement No. 107, which most entities have been employing for several years."

$$
\begin{aligned}
\frac{\Delta MVE_{it}}{MVE_{it-1}} = {} & \delta_0 + \sum_{t=1}^{T-1} \delta_{01,t} Y_t + \sum_{i=1}^{K-1} \delta_{02,i} BHC_i + \delta_{1,1} \frac{\Delta FVFA_{it}}{MVE_{it-1}} \\
& + \delta_{1,2} POST * \frac{\Delta FVFA_{it}}{MVE_{it-1}} + \delta_{2,1} \frac{\Delta FVFL_{it}}{MVE_{it-1}} + \delta_{2,2} POST * \frac{\Delta FVFL_{it}}{MVE_{it-1}} \\
& + \delta_{3,1} \frac{\Delta NFA_{it}}{MVE_{it-1}} + \delta_{3,2} POST * \frac{\Delta NFA_{it}}{MVE_{it-1}} + \delta_{4,1} \frac{\Delta NFL_{it}}{MVE_{it-1}} \\
& + \delta_{4,2} POST * \frac{\Delta NFL_{it}}{MVE_{it-1}} + \delta_5 \frac{\Delta DER_{it}}{MVE_{it-1}} + \delta_6 POST * \frac{\Delta DER_{it}}{MVE_{it-1}} \\
& + \delta_{7,1} \frac{\Delta NPL_{it}}{MVE_{it-1}} + \delta_{7,2} POST * \frac{\Delta NPL_{it}}{MVE_{it-1}} + \delta_{8,1} \frac{\Delta CORE_{it}}{MVE_{it-1}} \\
& + \delta_{8,2} POST * \frac{\Delta CORE_{it}}{MVE_{it-1}} + \varepsilon_{it}
\end{aligned}
\tag{4}
$$

where *DER* denotes net fair value of disclosed derivative financial instruments (*DNDP*) prior to the adoption of SFAS No. 133 and net fair value of recognized derivative financial instruments (*RNDP*) after the adoption of SFAS No. 133; *T* is the number of years in the sample period; Y_t equals 1 if the observation belongs to year *t*, and 0 otherwise; *K* is the number of bank holding companies in the sample; and *POST* equals 1 if the observation belongs to the post-SFAS No. 133 period, and 0 otherwise. In addition to *DER*, we also interact *FVFA*, *FVFL*, *NFA*, *NFL*, *NPL*, and *CORE* with *POST* to allow the coefficients on these variables to vary across pre- and post-SFAS No. 133 periods.

Note that moving from a disclosure regime to a recognition regime may affect the relative valuation of disclosed versus recognized derivatives through the effect of recognition on capital adequacy ratios.[10] However, these effects are unlikely to be significant if recognized derivatives are used to hedge assets or liabilities that are marked to market because any potential effects of recognized derivatives on Tier 1 capital will be offset by the losses or gains on the underlying assets or liabilities. Further, even if recognition results in increasing the volatility of capital ratios, the higher risk would bias against finding higher valuation coefficients for recognized items (Hodder et al. 2002).

IV. EMPIRICAL ANALYSIS

Sample Selection and Descriptive Statistics

The sample used for our tests comparing recognized and disclosed derivative instruments in the pre-SFAS No. 133 period consists of bank holding companies that (1) use derivatives for nontrading purposes between 1995 and 2000, (2) are publicly traded, (3) have 10-K reports in LexisNexis, and (4) have annual observations in the Bank Regulatory Database from 1995 to 2000.[11] The sample period is from 1995 to 2000 because both SFAS Nos. 107 and 119, which required expanded disclosures on fair values of derivative instruments and purposes of derivative usage, were effective during this period. This selection procedure yields a final sample of 146 bank holding companies. Of these, 58 bank holding companies have both recognized and disclosed derivative financial instruments held for risk management purposes, and 88 have only disclosed derivative financial instruments in the pre-SFAS No. 133 period.

---

[10] See the Joint Release entitled *"Interim Regulatory Reporting and Capital Guidance on SFAS No. 133, Accounting for Derivative Instruments and Hedging Activities"* issued by the OCC, Board of Governors of the Federal Reserve System, FDIC, and OTS dated December 29, 1998.

[11] The Bank Regulatory Database provides accounting data for bank holding companies, commercial banks, savings banks, and savings and loans institutions. The Bank Regulatory Database collects financial data from the Federal Reserve Bank of Chicago. The data is based on bank holding companies' FRY-9 reports.

December 31 stock price information for the bank holding companies is obtained from CRSP.[12] Fair value information for financial assets and liabilities including derivatives and notional value information for derivatives is collected from banks' 10-K reports in LexisNexis. We confirm these notional and fair value numbers with the FRY-9 reports in the Bank Regulatory Database.

Table 1 reports descriptive statistics for bank holding companies that have both recognized and disclosed risk management derivative financial instruments, referred to as the first group, and bank holding companies that only have disclosed derivatives, referred to as the second group.

A comparison of the first and the second group reveals that banks in the first group are considerably larger in terms of total assets than banks in the second group, and net derivative positions are economically significant at least for some banks. The net fair value of recognized derivatives ranges from −7.44 percent to 7.35 percent of book value of equity and net fair value of disclosed derivatives ranges from −11.02 percent to 14.59 percent of book value of equity for the first group, while net fair value of disclosed derivatives ranges from −18.31 percent to 23.76 percent of book value of equity for the second group. Further, for the first group, net fair value of disclosed derivatives ($57.35 million) is significantly higher, on average, than the net fair value of recognized derivatives ($3.24 million). The two groups of banks are similar in terms of the composition of their risk management derivative portfolio. For both groups, interest rate derivative contracts constitute the major portion of the overall risk management derivative portfolio. For the first group, in terms of notional amounts, interest rate derivative contracts constitute approximately 90 percent of the overall portfolio, while this figure is 91 percent for the second group. Descriptive statistics in Panel A of Table 1 also reveal that the compositions of the recognized and disclosed derivative instrument portfolios are similar. Interest rate contracts constitute 84.77 percent of the total notional amount of derivative contracts for recognized derivative instruments, and 90.07 percent for disclosed derivative instruments.

The sample for our second test, comparison of valuation of derivative instruments in the post-SFAS No. 133 period relative to valuation in the pre-SFAS No. 133 period, consists of 82 out of 88 bank holding companies, which have only disclosed derivative financial instruments before the adoption of SFAS No. 133 and which have the required data in the post-SFAS No. 133 period.[13] Our sample period is from 1995 to 2000 before the adoption of SFAS No. 133, and from 2001 to 2004 after the adoption. Because these 82 banks constitute approximately 93 percent of the second group of banks described in this section, their descriptive statistics are similar to those of the second group and, for this reason, are not reported separately.

When comparing valuation coefficients on derivative instruments before and after the adoption of SFAS No. 133, we need to rule out alternative explanations for the difference other than recognition versus disclosure. Considering that requirements of SFAS No. 133 have drastically changed the accounting for derivative instruments, there may have been a significant change in the extent of derivative usage or in the way derivatives are used.

---

[12] Using March 31 prices or returns cumulated to March 31 does not alter our inferences.

[13] Six bank holding companies are excluded from the pre- versus post-SFAS No. 133 analysis because they either merged with or were acquired by other bank holding companies.

TABLE 1
Descriptive Statistics

Panel A: Descriptive Statistics for Banks with Both Recognized and Disclosed Derivatives and Banks with Only Disclosed Derivatives

	Banks with Both Recognized and Disclosed Derivatives[a]				Banks with Only Disclosed Derivatives[b]			
Variables (Thousands)	**Mean**	**Median**	**Minimum**	**Maximum**	**Mean**	**Median**	**Minimum**	**Maximum**
CV of Financial Assets	68,348,191	34,455,477	1,149,562	834,300,000	10,338,256	4,000,152	213,751	213,818,000
FV of Financial Assets	69,025,566	34,778,203	1,163,840	850,800,000	10,864,600	4,089,322	216,027	214,652,000
CV of Financial Liabilities	63,582,641	31,566,062	1,009,786	773,200,000	9,680,893	3,685,743	199,485	186,959,000
FV of Financial Liabilities	63,662,633	31,803,545	1,017,434	773,600,000	9,688,872	3,701,404	199,405	186,293,000
Nonfinancial Assets	557,485	265,241	44,056	72,833,000	118,230	151,998	5,536	77,584,892
Nonfinancial Liabilities	432,946	234,602	19,064	70,506,137	90,482	100,467	3,276	71,056,562
RNDP	3,243	113,025	−326,000	692,000	NA	NA	NA	NA
DNDP	57,350	269,878	−1,068,000	1,426,000	2,830	5,304	−992,000	798,892
NPL	419,877	383,456	2,572	7,258,000	71,486	52,223	1,150	2,308,000
CORE	6,723,700	2,991,384	135,011	92,623,000	1,628,442	692,136	2,593	40,840,000
Net *FV* of Net Assets/*BVE*	0.1222	0.1252	0.0865	0.1664	0.1709	0.1483	0.1044	0.2523
RNDP/BVE	0.0029	0.0151	−0.0744	0.0735	NA	NA	NA	NA
DNDP/BVE	0.0275	0.0402	−0.1102	0.1459	0.0311	0.0385	−0.1831	0.2376
RNDP/MVE	0.0015	0.0088	−0.0521	0.0461	NA	NA	NA	NA
DNDP/MVE	0.0100	0.0112	−0.0592	0.1145	0.0124	0.0155	−0.1209	0.1108
NREC/NDER	0.2841	0.3022	0.0332	0.9267	NA	NA	NA	NA
NDISC/NDER	0.7159	0.6978	0.0733	0.9668	NA	NA	NA	NA
NRECIR/NREC	0.8477	0.7831	0.0000	1.000	NA	NA	NA	NA
NRECFX/NREC	0.1328	0.1221	0.0000	1.000	NA	NA	NA	NA
NRECEQ/NREC	0.0195	0.0000	0.0000	1.000	NA	NA	NA	NA
NDISCIR/NDISC	0.9007	0.8222	0.0000	1.000	0.9141	0.8777	0.0843	1.000
NDISCFX/NDISC	0.0975	0.1024	0.0000	1.000	0.0525	0.0608	0.0099	1.000
NDISCEQ/NDISC	0.0018	0.0135	0.0000	0.5393	0.0334	0.0252	0.0017	1.000

(*continued on next page*)

TABLE 1 (Continued)

Panel B: Banks Used in Pre- versus Post-SFAS No. 133 Analysis[c]

	Pre-SFAS No. 133 (1995–2000)	Post-SFAS No. 133 (2001–2004)
NDERIR/*NDER*	0.9334	0.9421
NDERFX/*NDER*	0.0337	0.0403
NDEREQ/*NDER*	0.0329	0.0176
NDER/Total Assets	0.0921	0.0888
NDER/*MVE*	0.5892	0.6308

---

[a] The first sample includes 342 bank-year observations for the 1995–2000 period (pre-SFAS No. 133 period) for 58 bank holding companies that have both recognized and disclosed derivatives.

[b] The second sample includes 528 bank-year observations for the 1995–2000 period (pre-SFAS No. 133 period) for 88 bank holding companies that have only disclosed derivatives.

[c] The sample includes 492 bank-year observations for the period 1995–2000 (pre-SFAS No. 133 period) and 328 bank-year observations for the period 2001–2004 (post- SFAS No. 133 period) for 82 bank holding companies.

Variable Definitions:

CV = carrying value;
FV = fair value;
RNDP = net fair value of recognized derivatives held for risk management purposes;
DNDP = net fair value of disclosed derivative instruments held for risk management purposes;
CORE = domestic deposits minus time deposits;
NPL = nonperforming loans;
BVE = book value of equity;
MVE = market value of equity;
NREC = total notional amount of recognized derivative instruments held for risk management purposes;
NDISC = total notional amount of disclosed derivative instruments held for risk management purposes;
NDER = total notional amount of derivative instruments held for risk management purposes (*NDER* = *NREC* + *NDISC*);
NRECIR = notional amount of recognized interest rate derivative instruments held for risk management purposes;
NRECFX = notional amount of recognized foreign-exchange derivative instruments held for risk management purposes;
NRECEQ = notional amount of recognized equity derivative instruments held for risk management purposes;
NDISCIR = notional amount of disclosed interest rate derivative instruments held for risk management purposes;
NDISCFX = notional amount of disclosed foreign-exchange derivative instruments held for risk management purposes;
NDISCEQ = notional amount of disclosed equity derivative instruments held for risk management purposes;
NDER = total notional amount of derivative instruments held for risk management purposes;
NDERIR = notional amount of interest rate derivative instruments held for risk management purposes;
NDERFX = notional amount of disclosed foreign-exchange derivative instruments held for risk management purposes; and
NDEREQ = notional amount of equity derivative instruments held for risk management purposes.

Panel B of Table 1 presents descriptive statistics regarding the extent of derivative usage and the types of derivatives used. Our tests reveal that there is no significant change in the extent of derivative usage and composition of the derivative instrument portfolio. In terms of notional amounts, interest rate derivatives constitute a major portion of the entire portfolio (93.34 percent and 94.21 percent of total notional amount of derivative contracts before and after SFAS No. 133, respectively). The descriptive statistics also suggest that there is no significant change in the extent of derivative usage after SFAS No. 133 relative to before SFAS No. 133. The total notional amount of derivative contracts equals 9.21 percent (58.92 percent) and 8.88 percent (63.08 percent) of total assets (market value of equity) before and after SFAS No. 133, respectively. These statistics suggest that any difference in the valuation of derivative financial instruments before and after SFAS No. 133 is unlikely to be attributable to change in the extent of derivative usage or composition of the derivative instrument portfolio.

Value-Relevance Tests of Recognition and Disclosure for Pre-SFAS No. 133 Period

Panel A of Table 2 presents estimation results for the cross-sectional levels model presented in Equation (1). These results are obtained using bank holding companies with both recognized and disclosed risk management derivative instruments. In order to control for heteroscedasticity and autocorrelation, we report t-statistics based on Newey and West's (1987) heteroscedasticity and autocorrelation corrected covariance estimates in the remaining tables.[14]

The results suggest that net fair value of recognized derivatives (*RNDP*) has significant explanatory power and is value-relevant. The coefficient on *RNDP*, α_5, is positive, as expected, and significantly different from zero (t-statistic = 3.03). The coefficient on *DNDP*, α_6, is not statistically different from zero at conventional levels (t-statistic = 0.91). This suggests that disclosed derivative fair values of banks have no power in explaining cross-sectional variation in banks' market values of equity. The coefficients on other balance sheet variables (*FVFA*, *FVFL*, *NFA*, and *NFL*) have the predicted signs and are significantly different from zero. The coefficients on the control variables (*NPL* and *CORE*) also have the predicted signs though only the coefficient on *CORE* is reliably different from zero. We also test whether the coefficient on *RNDP*, α_5, is statistically greater than the coefficient on *DNDP*, α_6. Our findings suggest that α_5 is greater than α_6 at the 1 percent level, suggesting that the market places a lower weight on disclosed derivatives.

Panel B of Table 2 presents the results for the changes model for bank holding companies that have both recognized and disclosed derivatives. The results of the changes model are not discussed in detail because the changes model provides results very similar to the levels model.

A potential explanation for not finding a significant coefficient on *DNDP* is that the tests lack power because of the relatively small sample size. To assess this, we use a combined sample that includes banks with both recognized and disclosed derivatives and banks with only disclosed derivatives prior to SFAS No. 133 and test the significance of the valuation coefficients. In untabulated results, consistent with the results in Table 2, we find that the valuation coefficients on recognized derivatives are significant, whereas the valuation coefficients on disclosed derivatives are not significant.

---

[14] We use the Newey-West fixed bandwidth, q, in our tests (Newey and West 1987). It is based on the number of observations in the sample, and is computed as: $q = \text{int}\,[4(N/100)^{2/9}]$, where int [ ] denotes the integer part and N is the number of observations.

TABLE 2
Tests of Value-Relevance of Recognized versus Disclosed Derivative Instrument Fair Values in the Pre-SFAS No. 133 Period for the Sample of Banks with Both Recognized and Disclosed Derivative Instruments

Panel A: Levels Model[a]

Variable[b,c]	Predicted Sign	Coefficient Estimates	t-statistic[d]	p-value[e]
FVFA	+	1.379	4.321	0.001
FVFL	–	−1.375	−4.219	0.001
NFA	+	1.338	3.996	0.001
NFL	–	−1.293	−4.103	0.001
RNDP	+	1.270	3.025	0.001
DNDP	+	0.635	0.905	0.183
NPL	–	−0.459	−0.781	0.218
CORE	+	0.476	5.086	0.001
Adj. R^2		0.882		
No. of Obs.		342		

Panel B: Changes Model[f]

Variable[g,h]	Predicted Sign	Coefficient Estimates	t-statistic[c]	p-value[d]
Δ*FVFA*	+	0.678	3.621	0.001
Δ*FVFL*	–	−0.667	−3.333	0.001
Δ*NFA*	+	0.686	3.801	0.001
Δ*NFL*	–	−0.668	−3.603	0.001
Δ*RNDP*	+	0.854	2.661	0.004
Δ*DNDP*	+	0.528	0.896	0.174
Δ*NPL*	–	−0.462	−2.223	0.014
Δ*CORE*	+	0.340	1.417	0.079
Adjusted R^2		0.375		
No. of Observations		284		

---

[a] The sample includes annual observations from 1995 to 2000 for banks that have both recognized and disclosed derivatives.

[b] For bank i at time t, MVE_{it} = market value of equity; $FVFA_{it}$ = aggregate fair value of financial assets excluding risk management derivative instruments; $FVFL_{it}$ = aggregate fair value of financial liabilities excluding risk management derivative instruments; NFA_{it} = aggregate book value of nonfinancial assets; NFL_{it} = aggregate book value of nonfinancial liabilities; $RNDP_{it}$ = net fair value of recognized derivative instruments; $DNDP_{it}$ = net fair value of disclosed derivative instruments; NPL_{it} = nonperforming loans; $CORE_{it}$ = domestic deposits minus time deposits; CSO_{it} = number of common shares outstanding; $Y_t = 1$ if the observation belongs to year $t = T$ for $T = 1, 2, \ldots, 6$, and 0 otherwise; and $BHC_i = 1$ if the observation belongs to bank holding company $i = K$ for $K = 1, 2, \ldots, 58$, and 0 otherwise.

[c] The F-statistic for testing the hypothesis that $\alpha_{01,1} = \alpha_{01,2} = \ldots \alpha_{01,5} = 0$ and the F-statistic for testing the hypothesis that $\alpha_{02,1} = \alpha_{02,2} = \ldots \alpha_{02,57} = 0$ indicate that the null hypothesis can be rejected at the 1 percent level.

[d] All t-statistics are based on Newey and West's (1987) heteroscedasticity and autocorrelation consistent standard error estimates.

[e] All p-values are based on one-tailed t-tests when the coefficient sign is predicted, and based on two-tailed t-tests otherwise.

(*continued on next page*)

TABLE 2 (Continued)

[f] The sample includes annual observations from 1995 to 2000 for banks that have both recognized and disclosed derivatives.

[g] For bank *i* during year *t*, ΔMVE_{it} = change in market value of equity; $\Delta FVFA_{it}$ = change in aggregate fair value of financial assets excluding risk management derivative instruments; $\Delta FVFL_{it}$ = change in aggregate fair value of financial liabilities excluding risk management derivative instruments; $\Delta RNDP_{it}$ = change in net fair value of recognized derivative instruments; $\Delta DNDP_{it}$ = change in net fair value of disclosed derivative instruments; ΔNPL_{it} = change in nonperforming loans; $\Delta CORE_{it}$ = change in domestic deposits minus time deposits; $Y_t = 1$ if the observation belongs to year $t = T$ for $T = 1, 2, \ldots, 5$, and 0 otherwise; and $BHC_i = 1$ if the observation belongs to bank holding company $i = K$ for $K = 1, 2, \ldots, 58$, and 0 otherwise.

[h] The F-statistic for testing the hypothesis that $\beta_{01,1} = \beta_{01,2} = \ldots = \beta_{01,4} = 0$ and the F-statistic for testing the hypothesis that $\beta_{02,1} = \beta_{02,2} = \ldots = \beta_{02,57} = 0$ indicate that the null hypothesis can be rejected at the 1 percent level.

To summarize, our results indicate that, in the pre-SFAS No. 133 period, recognized fair values of risk management derivative instruments are reliably relevant in explaining equity valuation of bank holding companies, whereas disclosed fair values of risk management derivative instruments are not.[15]

Value-Relevance Tests of Recognition and Disclosure for Pre- versus Post-SFAS No. 133 Periods

The analysis in this section uses the banks that have only disclosed derivatives in the pre-SFAS No. 133 period. Panel A of Table 3 presents estimation results of the levels model presented in Equation (3). These results suggest that, for the 82 bank holding companies in our sample, the coefficient on net fair value of derivative financial instruments disclosed prior to SFAS No. 133 (λ_5) is not significantly related to market value of equity (t-statistic = 0.86). On the other hand, net fair value of derivative financial instruments recognized after SFAS No. 133 is significantly positively related to market value of equity (t-statistic for $\lambda_5 + \lambda_6 = 2.18$). The positive and significant coefficient on *POST*DER* (t-statistic = 4.93), λ_6, suggests that derivative fair values, which were disclosed prior to SFAS No. 133, are given more weight in valuation after SFAS No. 133, which mandates their recognition.

The coefficients on the other balance sheet and control variables exhibit the predicted signs and are significant at conventional levels. The coefficients on interaction terms, except for *POST*DER*, are not reliably different from zero, suggesting that other balance sheet and control variables are given the same weight in valuation in both the pre- and the post-SFAS No. 133 periods.

The estimation results for the changes models, presented in Panel B of Table 3, are consistent with the results for the levels models. They show that the coefficient on changes in net fair values of derivative financial instruments disclosed prior to SFAS No. 133, δ_5, is not significant in explaining changes in banks' equity values (t-statistic = 0.45), whereas the coefficient on changes in net fair values of derivative financial instruments recognized after SFAS No. 133, $\delta_5 + \delta_6$, is significant (t-statistic = 1.92). The coefficient on *POST*ΔDER*, λ_6, is positive and statistically significant (t-statistic = 3.11), suggesting that the market places more weight on derivative fair values when they are recognized.

---

[15] These findings are in contrast to Venkatachalam (1996), who finds that derivative fair value disclosures are value relevant, but are consistent with Barth et al. (1996), Eccher et al. (1996), and Nelson (1996), who report that fair values of banks' off-balance-sheet derivative financial instruments are not significantly related to equity values.

TABLE 3
Tests of Value-Relevance of Derivative Instrument Fair Values in the Post- and Pre-SFAS No. 133 Period for the Sample of Banks with Only Disclosed Derivative Financial Instruments in the Pre-SFAS No. 133 Period

Panel A: Levels Model[a]

Variable[b,c]	Predicted Sign	Coefficient Estimates	t-statistic[d]	p-value[e]
FVFA	+	1.266	5.444	0.001
POST*FVFA	?	−0.054	−0.193	0.847
FVFL	−	−1.238	−5.212	0.001
POST*FVFL	?	0.026	0.101	0.920
NFA	+	1.004	5.224	0.001
POST*NFA	?	−0.007	−0.910	0.363
NFL	−	−0.961	−5.303	0.001
POST*NFL	?	0.004	1.223	0.222
DER	+	0.673	0.863	0.194
POST*DER	?	0.787	4.928	0.001
NPL	−	−0.383	−4.676	0.001
POST*NPL	?	0.179	1.263	0.207
CORE	+	0.135	2.359	0.009
POST*CORE	?	−0.046	−0.402	0.688
Adjusted R^2		0.853		
No. of Obs.		820		
t ($\lambda_5 + \lambda_6$)	+		2.181	0.015

Panel B: Changes Model[f]

Variable[g,h]	Predicted Sign	Coefficient Estimates	t-statistic[d]	p-value[e]
ΔFVFA	+	0.618	3.099	0.001
POST*ΔFVFA	?	−0.134	−0.936	0.350
ΔFVFL	−	−0.579	−3.802	0.001
POST*ΔFVFL	?	0.149	0.888	0.375
ΔNFA	+	0.754	2.554	0.005
POST*ΔNFA	?	−0.076	−2.718	0.007
ΔNFL	−	−0.700	−2.301	0.011
POST*ΔNFL	?	0.057	2.277	0.023
ΔDER	+	0.256	0.453	0.325
POST*ΔDER	?	0.534	3.112	0.002
ΔNPL	−	−0.110	−0.850	0.198
POST*ΔNPL	?	0.079	0.776	0.438
ΔCORE	+	0.121	1.320	0.094
POST*ΔCORE	?	0.153	0.691	0.490
Adjusted R^2		0.411		
No. of Obs.		656		
t ($\delta_5 + \delta_6$)	+		1.920	0.028

(continued on next page)

TABLE 3 (Continued)

[a] The sample includes annual observations from 1995 to 2004.

[b] For bank i at time t, MVE_{it} = market value of equity; $FVFA_{it}$ = aggregate fair value of financial assets excluding risk management derivative instruments; $FVFL_{it}$ = aggregate fair value of financial liabilities excluding risk management derivative instruments; NFA_{it} = aggregate book value of nonfinancial assets; NFL_{it} = aggregate book value of nonfinancial liabilities; DER_{it} = net fair value of disclosed derivative instruments if the observation belongs to the pre-SFAS No. 133 period, or = net fair value of recognized derivative instruments if the observation belongs to the post-SFAS No. 133 period; NPL_{it} = nonperforming loans; $CORE_{it}$ = domestic deposits minus time deposits; CSO_{it} = number of common shares outstanding; Y_t = 1 if the observation belongs to year $t = T$ for $T = 1, 2, \ldots, 9$, and 0 otherwise; BHC_i = 1 if the observation belongs to bank holding company $i = K$ for $K = 1, 2, \ldots, 82$, and 0 otherwise; and $POST$ = 1 if the observation belongs to post-SFAS No. 133 period, and 0 otherwise.

[c] The F-statistic for testing the hypothesis that $\lambda_{01,1} = \lambda_{01,2} = \ldots \lambda_{01,9} = 0$ and the F-statistic for testing the hypothesis that $\lambda_{02,1} = \lambda_{02,2} = \ldots \lambda_{02,81} = 0$ indicates that the null hypothesis can be rejected at the 1 percent level.

[d] All t-statistics are based on Newey and West's (1987) heteroscedasticity and autocorrelation consistent standard error estimates.

[e] All p-values are based on one-tailed t-tests when the coefficient sign is predicted, and based on two-tailed t-tests otherwise.

[f] The sample includes annual observations from 1995 to 2004.

[g] Variable definitions: For bank i during year t, ΔMVE_{it} = change in market value of equity; $\Delta FVFA_{it}$ = change in aggregate fair value of financial assets excluding risk management derivative instruments; $\Delta FVFL_{it}$ = change in aggregate fair value of financial liabilities excluding risk management derivative instruments, or ΔDER_{it} = change in net fair value of disclosed derivative instruments if the observation belongs to the pre-SFAS No. 133 period; = change in net fair value of recognized derivative instruments if the observation belongs to the post-SFAS No. 133 period; ΔNPL_{it} = change in nonperforming loans; $\Delta CORE_{it}$ = change in domestic deposits minus time deposits; Y_t = 1 if the observation belongs to year $t = T$ for $T = 1, 2, \ldots, 7$, and 0 otherwise; BHC_i = 1 if the observation belongs to bank holding company $i = K$ for $K = 1, 2, \ldots, 82$, and 0 otherwise; and $POST$ = 1 if the observation belongs to post-SFAS No. 133 period, and 0 otherwise.

[h] The F-statistic for testing the hypothesis that $\delta_{01,1} = \delta_{01,2} = \ldots = \delta_{01,7} = 0$ and the F-statistic for testing the hypothesis that $\delta_{02,1} = \delta_{02,2} = \ldots = \delta_{02,81} = 0$ indicate that the null hypothesis can be rejected at the 1 percent level.

The coefficients on the other balance sheet variables have the predicted signs and are significant at conventional levels. The coefficients on two interaction terms, $POST{*}\Delta NFA$ and $POST{*}\Delta NFL$, are significant, suggesting that nonfinancial balance sheet accounts are not given the same weight in valuation in the pre- and the post-SFAS No. 133 periods. Unlike the results for the levels specification, the coefficients on the control variables, ΔNFL and $\Delta CORE$, are insignificant (t-statistic = −0.85), and marginally significant (t-statistic = 1.32), respectively.

Our analysis in this section may be difficult to interpret if banks change their risk management strategies significantly with the adoption of SFAS No. 133. Specifically, a change, from the pre-SFAS No. 133 to the post-SFAS No. 133 period, in how risk management derivatives are used might cloud our findings. While there has been speculation on whether and how SFAS No. 133 may affect firms' risk management strategies, empirical evidence shows no significant effect on banks.[16] According to a survey by the Association for Financial Professionals (AFP 2002) over 70 percent of the respondents state that SFAS No. 133 has had no impact on the way they use derivative instruments. Results from a similar survey conducted by Treasury & Risk Management (2001) also indicate that almost

---

[16] For example, in his article in *CFO Magazine*, Radigan (2001) says, "banking companies that already have rigorous accounting standards in place are finding life under SFAS No. 133 may not be that different from what came before."

70 percent of the derivative-users stated that their derivatives strategies have not changed in light of SFAS No. 133.

We conduct two tests to examine whether banks changed their hedging strategies as they moved from a disclosure environment to a recognition environment. For derivatives held for risk management purposes, gains and losses on these contracts are expected to be negatively associated with gains and losses on hedged assets and liabilities. Following Venkatachalam (1996), we test the existence of such a relation for the sample of banks used in the pre- versus post-SFAS No. 133 analysis. We analyze the correlation coefficients between fair value gains and losses on risk management derivatives (*GLDER*) and fair value gains and losses on-balance-sheet assets and liabilities (*GLBS*). Our definition of on-balance-sheet assets and liabilities includes investment securities, net loans, deposits, and long-term debt. These financial assets and liabilities are inherently exposed to interest rate and exchange rate risk, thus constituting most commonly hedged balance sheet items (Venkatachalam 1996). Fair value gains and losses are computed as the sum of realized gains and losses and changes in unrealized gains and losses (Barth 1994; Venkatachalam 1996). Our analysis (not reported) reveals that the correlations between *GLDER* and *GLBS* are -0.33 and -0.30 in the pre-SFAS No. 133 and the post-SFAS No. 133 periods, respectively. The negative correlations suggest that the derivatives we examine are, in aggregate, used for offsetting risk exposures inherent in on-balance-sheet financial assets and liabilities both prior to and after SFAS No. 133.

In addition to analyzing the direction of the relationship, we also test whether the magnitude of the correlation changed from the pre-SFAS No. 133 to the post-SFAS No. 133 period. Using the procedure developed by Fisher (1921), we test whether the correlation between *GLDER* and *GLBS* during the pre-SFAS No. 133 period is different from the correlation during the post-SFAS No. 133 period. We find that the correlations are not different (z-statistic $= -0.44$, and p-value $= 0.66$) across the two periods, suggesting that the banks in our sample are unlikely to have changed their hedging strategies following SFAS No. 133.

Self-Selection Correction

A potential problem with our tests based on the sample of banks that held both recognized and disclosed derivatives or banks that held only disclosed derivatives pre-SFAS No. 133 is that banks could self-select into either group. We control for this self-selection bias using the two-stage estimation procedure developed by Heckman (1976, 1979) and Lee (1976, 1978). We estimate the following probit model to explain the classification of banks into the two groups:

$$\begin{aligned} R_{it} = {} & \psi_1 + \psi_2 AFS_{it} + \psi_3 SIZE_{it} + \psi_4 COD_{it} + \psi_5 CAP_{it} + \psi_6 NIM_{it} \\ & + \psi_7 IRREXP_{it} + \varepsilon_{it} \end{aligned} \tag{5}$$

where R_{it} equals 1 if bank i holds risk management derivative financial instruments recognized at fair value, and 0 otherwise; AFS_{it} is fair value of available-for-sale securities scaled by total assets; $SIZE_{it}$ is natural logarithm of total assets; COD_{it} is interest expense on short- and long-term debt, subordinated notes, and debentures scaled by the sum of debt, subordinated notes, and debentures; CAP_{it} is equity capital scaled by total assets; NIM_{it} is net interest income scaled by total assets; and $IRREXP_{it}$ is one-year maturity gap scaled by total assets. These variables proxy for the likelihood of banks holding recognized derivatives.[17] The probit model classifies 82.46 percent (79.92 percent) of the observations correctly for the first (second) group, indicating that the probit model has high classificatory power.[18] In untabulated results, we find that including the inverse-Mills ratio in our levels and changes models does not alter any of our inferences.

Additional Sensitivity Analyses

We perform a number of additional tests and sensitivity analyses. First, we repeat our tests using annual year-by-year regressions. Although we only have a small number of annual observations, we find the coefficient on *RNDP* to be statistically significant in five of the six years (1995, 1996, 1997, 1999, 2000), and the coefficient on *DNDP* to be insignificant in all six years. Second, to reduce the potential bias in coefficient standard error estimates that may result from cross-sectional correlation, we compute a t-statistic from the regression coefficient estimates obtained from annual regressions (Bernard 1987). Using this alternative t-statistic does not alter any of our inferences. Third, for both the pre-SFAS No. 133 and the pre- versus post-SFAS No. 133 analysis, we examine the sensitivity of our results to the choice of deflator by repeating levels estimations using variables scaled by beginning book value of equity in place of number of shares outstanding. We obtain similar results as those reported except that the coefficient on recognized derivative fair values following SFAS No. 133, $\lambda_5 + \lambda_6$, is also not different from zero. However, we find that recognized derivatives are significantly valued when the post-SFAS No. 133 period is defined as the period 2002–2004. Finally, to rule out the possibility that our pre- versus post-SFAS No. 133 results are driven by an increase in value-relevance of recognized derivatives for some reason other than recognition versus disclosure, we conducted a pre- versus post-SFAS No. 133 analysis for the banks that had both recognized and disclosed derivatives prior to SFAS No. 133. Our results (not reported) show that value-relevance of recognized derivatives has not changed from the pre- to the post-SFAS No. 133 period.

Possible Differences in Reliability

While our findings are consistent with recent limited attention or costly information-processing models, an alternative explanation for our findings is that there are differences in the reliability of recognized and disclosed derivatives. However, two factors might mitigate the reliability differences between recognized and disclosed fair values of derivatives.

First, all derivative contracts are traded either on exchanges, or in over-the-counter (OTC) markets. For exchange-traded derivatives, price verification is straightforward as they are traded in highly liquid markets. Although OTC markets are not as liquid as exchanges, they operate in a very similar manner to exchanges. Dodd (2002) notes that in all three types of OTC markets (traditional dealer markets, electronically brokered markets, and proprietary electronic dealer platforms), the trading process is either bilateral or multilateral, and the ability to instantaneously communicate among dealers and customers enables market participants to execute trades in a multilateral environment similar to exchanges as defined by the Commodity Exchange Act. Smith and Wilson (1996) also state that interest rate swaps are traded enough to obtain quotes from dealers. For these reasons, although not as liquid as exchanges, OTC markets are still considered liquid for derivative products such as interest rate swaps.[19]

---

[17] We expect the probability of holding recognized derivatives to be increasing with the amount of available-for-sale securities, bank size (Angbazo 1997), cost of external financing (Froot and Stein 1998), inherent interest rate risk exposure (Carter and Sinkey 2000; Gunther and Siems 2002; Reichert and Shyu 2002), and decreasing with capital adequacy ratio and bank performance (Ahmed et al. 1999).

[18] Consistent with prior studies (e.g., Chaney et al. 2004), we assess the classificatory power of the first stage probit model using a 0.50 cutoff.

Second, observable market inputs are the major determinants of the fair values of derivatives traded in the OTC market. The Public Company Accounting Oversight Board (2004) states that verification of reliability of a fair value measurement depends on the extent to which market-based inputs, which are easily verified, and entity-specific assumptions, which are the results of unobservable management expectations, are used in calculations. Fair values of derivative contracts including interest rate swaps and foreign exchange forwards, which constitute over 90 percent of banks' derivatives portfolios (Office of the Comptroller of the Currency 2003), depend mostly on market-based inputs including rates, prices, and indices to generate continuous yield or pricing curves. The pricing of these derivatives also includes entity-specific factors, mainly counter-party default risk, to some extent. However, entity-specific information plays a relatively smaller role in pricing derivatives (Litzenberger 1992).

While these factors mitigate reliability differences between recognized and disclosed derivative fair values, we cannot definitively rule out differences in perceived reliability as an alternative explanation because there is some evidence that auditors might impose higher standards for recognized information than for disclosed information (Libby et al. 2005).

V. CONCLUSION

This study investigates the value-relevance of banks' derivative financial instruments held for risk management purposes that are recognized in financial statements at fair value and those that are not recognized, but whose fair values are disclosed in notes to the financial statements. Expanded disclosures and accounting practices prior to and after SFAS No. 133 provide a natural setting for comparing the valuation implications of recognized and disclosed derivative fair value information.

Using a sample of banks that simultaneously hold recognized and disclosed derivatives prior to SFAS No. 133, we find that valuation coefficients on the recognized derivatives are significantly positive, whereas valuation coefficients on the disclosed derivatives are not significantly different from zero. Further, using a sample of banks that have only disclosed derivatives prior to SFAS No. 133, which are recognized after SFAS No. 133, we find that while the valuation coefficients on disclosed derivatives were not significant, the valuation coefficients on recognized derivatives (after SFAS No. 133) are significantly positive.

Our findings suggest that SFAS No. 133 has been successful in increasing the transparency and visibility of derivative financial instruments. Our evidence is also consistent with conjectures of practitioners as well as recent models of limited attention or costly information processing that suggest that investors do not pay an equal amount of attention to disclosed amounts relative to recognized amounts. Reliability differences between recognized and disclosed derivative fair values might be an alternative explanation for our findings.

Our results are in contrast to the prevalent implicit assumption in empirical accounting research that whether an item is recognized or disclosed does not matter in terms of value-relevance and suggest that recognition and disclosure are not substitutes. Finally, our results also suggest that researchers should control for potentially different pricing implications of recognized and disclosed information in their studies.

---

[19] Recent empirical studies in finance also show that the OTC derivatives markets are informationally efficient. For example, Gupta and Subrahmanyam (2000) document evidence of mispricing of swaps during the 1980s and early 1990s, and find that the mispricing is eliminated with increased liquidity starting from the mid-1990s.

APPENDIX
Disclosures from Bank Annual Reports about How Banks Define Their Trading and Non-Trading Derivative Activities

From the 2000 Annual Report of Keycorp, Inc.:

Key engages in capital market, primarily through the Key Capital Partners line of business. These activities encompass a variety of services. Among other things, we trade securities as a **dealer**, enter into derivative contracts (both to **accommodate clients' financing needs** and **for proprietary trading purposes**), invest in privately held companies and conduct transactions in foreign currencies (both to accommodate clients' needs and to benefit from fluctuations in exchange rates).

DERIVATIVES USED FOR ASSET AND LIABILITY MANAGEMENT PURPOSES

Key uses derivatives known as interest rate swaps and caps to manage interest rate risk. These instruments **modify the repricing or maturity characteristics of specified on-balance-sheet assets and liabilities.** For example, an interest rate cap tied to variable rate debt would effectively prevent the interest rate on that debt from rising above a specified point.

(paragraph omitted)

DERIVATIVES USED FOR TRADING PURPOSES

Derivatives that are not used for asset and liability management purposes are used for **trading purposes**. Key enters into contracts for such derivatives either to **make a market for clients** or **for proprietary trading purposes**. Derivatives used for trading purposes typically include financial futures, foreign exchange forward and spot contracts, written and purchased options (including currency options), and interest rate swaps, caps, and floors. All derivatives used for trading purposes are recorded at fair value.

From the 2000 Annual Report of Citigroup, Inc.:

Citigroup enters into derivative and foreign exchange futures, forwards, options, and swaps, which **enable customers to transfer, modify, or reduce their interest rate, foreign exchange, and other market risks**, and **also trades these products for its own account**. In addition, Citigroup uses derivatives and other instruments, primarily interest rate products, as **an end-user** in connection with its **risk management activities.** Derivatives are used to manage interest rate risk relating to specific groups of on-balance-sheet assets and liabilities, including investments, commercial and consumer loans, deposit liabilities, long-term debt, and other interest-sensitive assets and liabilities, as well as credit card securitizations, redemptions, and sales. In addition, foreign exchange contracts are used to hedge non-U.S. dollar denominated debt, net capital exposures, and foreign exchange transactions.

(paragraph omitted)

The following table provides data on the notional principal amounts and maturities of **end-user (non-trading)** derivatives, along with additional data on end-user interest rate swaps and net purchased option positions at year-end 2000 with three-month LIBOR forward rates included for reference.

REFERENCES

Aboody, D. 1996. Recognition versus disclosure in the oil and gas industry. *Journal of Accounting Research* 34 (Supplement): 21–32.

Ahmed, A. S., and C. Takeda. 1995. Stock market valuation of gains and losses on commercial banks' investment securities. *Journal of Accounting and Economics* 20 (September): 207–225.

———, A. Beatty, and B. Bettinghaus. 1999. Interest rate risk management of bank holding companies: An examination of trade-offs in the use of investment securities and interest rate swaps. Working paper, The Pennsylvania State University.

American Institute of Certified Public Accountants (AICPA). 2000. *Audit and Accounting Guide for Banks and Savings Institutions*. New York, NY: AICPA.

Angbazo, L. 1997. Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking. *Journal of Banking and Finance* 21 (January): 55–87.

Association for Financial Professionals (AFP). 2002. *The Impact of SFAS No. 133 on the Risk Management Practices of End Users of Derivatives*. Bethesda, MD: AFP.

Barth, M. E. 1994. Fair value accounting: Evidence from investment securities and the market valuation of banks. *The Accounting Review* 69 (January): 1–25.

———, W. H. Beaver, and W. R. Landsman. 1996. Value-relevance of banks' fair value disclosures under SFAS No. 107. *The Accounting Review* 71 (October): 513–537.

———, G. J. Clinch, and T. Shibano. 2003. Market effects of recognition and disclosure. *Journal of Accounting Research* 41 (September): 581–609.

Bernard, V. L. 1987. Cross-sectional dependence and problems in inference in market-based accounting research. *Journal of Accounting Research* 25 (Spring): 1–48.

———, and K. Schipper. 1994. Recognition and disclosure in financial reporting. Working paper, University of Michigan and University of Chicago.

Carter, D. A., and J. F. Sinkey. 2000. Evidence on the financial characteristics of banks that do and do not use derivatives. *Quarterly Review of Economics and Finance* 40 (Winter): 431–449.

Chaney, P. K., D. C. Jeter, and L. Shivakumar. 2004. Self-selection of auditors and audit pricing in private firms. *The Accounting Review* 79 (January): 51–72.

Davis-Friday, P. Y., C. Liu, and H. F. Mittelstaedt. 2004. Recognition and disclosure reliability: Evidence from SFAS No. 106. *Contemporary Accounting Research* 21 (Summer): 399–427.

Dodd, R. 2002. The structure of OTC derivatives markets. *The Financier* 9 (Spring–Winter): 41–45.

Eccher, A., K. Ramesh, and S. R. Thiagarajan. 1996. Fair value disclosures of bank holding companies. *Journal of Accounting and Economics* 22 (August–December): 79–117.

Emerging Issues Task Force (EITF). 1984a. *Termination of Interest Rate Swaps*. Issue 84-07. Norwalk, CT: FASB.

——— 1984b. *Interest Rate Swap Transactions*. Issue 84-36. Norwalk, CT: FASB.

———. 1988. *Mortgage Swaps*. Issue 88-08. Norwalk, CT: FASB.

———. 1990. *Hedging Foreign Risk with Purchased Options Contracts*. Issue 90-17. Norwalk, CT: FASB.

———. 1991. *Hedging Foreign Currency Risks with Complex Options*. Issue 91-04. Norwalk, CT: FASB.

Financial Accounting Standards Board (FASB). 1981. *Foreign Currency Translation*. Statement of Financial Accounting Standards No. 52. Stamford, CT: FASB.

———. 1984. *Accounting for Future Contracts*. Statement of Financial Accounting Standards No. 80. Stamford, CT: FASB.

———. 1990. *Disclosure of Information about Financial Instruments with Off-Balance-Sheet Risk and Financial Instruments with Concentrations of Credit Risk*. Statement of Financial Accounting Standards No. 105. Stamford, CT: FASB.

———. 1991. *Disclosures about Fair Values of Financial Instruments*. Statement of Financial Accounting Standards No. 107. Stamford, CT: FASB.

———. 1994. *Disclosure about Derivative Financial Instruments and Fair Value of Financial Instruments*. Statement of Financial Accounting Standards No. 119. Stamford, CT: FASB.

———. 1998. *Accounting for Derivative Instruments and Hedging Activities*. Statement of Financial Accounting Standards No. 133. Stamford, CT: FASB.

Flannery, M. 1982. Retail bank deposits as quasi-fixed factors of production. *American Economic Review* 72 (June): 527–536.

———, and C. James. 1984. The effect of interest rate changes on the common stock changes of financial institutions. *Journal of Finance* 39 (September): 1141–1153.

Froot, K. A., and J. C. Stein. 1998. Risk management, capital budgeting, and capital structure policy for financial institutions: An integrated approach. *Journal of Financial Economics* 47 (January): 55–83.

Gunther, J. W., and T. F. Siems. 2002. The likelihood and extent of banks' involvement with interest rate derivatives as end users. *Research in Finance* 19: 125–142.

Gupta, A., and M. G. Subrahmanyam. 2000. An empirical examination of the convexity bias in the pricing of interest rate swaps. *Journal of Financial Economics* 55 (February): 239–279.

Heckman, J. 1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5: 475–492.

———. 1979. The sample selection bias as a specification error. *Econometrica* 47 (January): 153–162.

Hirshleifer, D. A., and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36 (December): 337–387.

Hodder, L., M. Kohlbeck, and M. L. McAnally. 2002. Accounting choices and risk management: SFAS No. 115 and U.S. bank holding companies. *Contemporary Accounting Research* 19 (Summer): 225–270.

Kothari, S. P., and J. L. Zimmerman. 1995. Price and return models. *Journal of Accounting and Economics* 20 (September): 155–193.

Lee, L. F. 1976. Estimation of limited dependent variable models by two-stage methods. Ph.D. dissertation, University of Rochester.

———. 1978. Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *International Economic Review* 19 (June): 415–433.

Libby, R., M. Nelson, and J. Hunton. 2005. Recognition vs. disclosure and auditor misstatement correction: The cases of stock compensation and leases. Working paper, Cornell University.

Litzenberger, R. 1992. Swaps: Plain and fanciful. *Journal of Finance* 47 (July): 831–850.

Nelson, K. 1996. Fair value accounting for commercial banks: An empirical analysis of SFAS No. 107. *The Accounting Review* 71 (April): 161–183.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (May): 703–708.

Office of the Comptroller of the Currency (OCC). 2003. *Bank Derivatives Report*. First quarter. Washington, D.C.: OCC.

Public Company Accounting Oversight Board (PCAOB). 2004. *Auditing Fair Value*. Washington, D.C.: PCAOB.

Radigan, J. 2001. *FAS 133 and the New Facts of Life*. *CFO Magazine* (March 7).

Reichert, A. K., and Y. Shyu. 2002. The determinants of derivative use by U.S. and foreign banks. *Research in Finance* 19: 143–172.

Smith, G. R., and A. C. Wilson. 1996. Proposed accounting for interest rate swaps. *CPA Journal* 66 (October): 54–58.

Treasury & Risk Management. 2001. *2001 Derivatives Survey*. New York, NY: Treasury & Risk Management.

Venkatachalam, M. 1996. Value relevance of banks' derivative disclosures. *Journal of Accounting and Economics* 22 (August–December): 327–355.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 589–616

The Persistence, Forecasting, and Valuation Implications of the Tax Change Component of Earnings

Andrew P. Schmidt
Columbia University

ABSTRACT: I examine whether earnings generated by changes in effective tax rates (the tax change component) persist and aid in forecasting future earnings. In addition, this study investigates to what extent investors incorporate the forecasting implications of the tax change component of earnings into stock prices. I find that there is a positive, significant association between the tax change component of earnings and future earnings. I use the interim reporting requirements of APB No. 28 (APB 1973) and FASB Interpretation No. 18 (FASB 1977) to further decompose the tax change component into an initial and a revised portion based on the first quarter estimate of the annual effective tax rates (ETR). I find that the initial tax change component is more persistent for future earnings than the revised tax change component. These results are consistent with my hypotheses that the initial and revised tax change components have differential persistence and forecasting implications, and dispute the broad notion advanced by prior literature that ETR-related earnings changes are transitory. Results from market tests indicate that the market underweights the forecasting implications of the tax change component and the mispricing appears to be driven by the transitory nature of the revised tax change component.

Keywords: *market efficiency; effective tax rates; earnings persistence; valuation.*

JEL Classification: *G14; H25; M41.*

I appreciate helpful comments and suggestions from my dissertation committee: Charles Christian (Co-Chair), Joseph Comprix (Co-Chair), James Boatsman, Lillian Mills (The University of Arizona), and Sanjay Gupta, as well as those of T. J. Atwood, Michael Hertzel, Mary Ann Hofmann, Steve Kaplan, James McKeown, Jeff Miller, Tom Omer, Steven Penman, John Philips, Philip Regier, Sonja Rego, Janet Samuels, Terry Shevlin (editor), Steven Smith, David Weber, Michelle Yetman, two anonymous reviewers, and workshop participants at Columbia University, The Pennsylvania State University, Texas A&M University, University of Illinois at Urbana–Champaign, The University of Iowa, University of Missouri, University of Notre Dame, Tulane University, the 2003 Illinois Tax Research Symposium, and the 2003 American Accounting Association Annual Meeting. Financial assistance from the Arizona State University Graduate Research Support Program is gratefully acknowledged. Special thanks to Wayne Thomas for generously providing code for and assistance with the Mishkin (1983) procedure. This paper won the 2005 ATA PricewaterhouseCoopers Best Tax Dissertation Award.

Editor's note: This paper was accepted by Terry Shevlin.

Submitted November 2003
Accepted October 2005

I. INTRODUCTION

This study investigates whether earnings generated by changes in the effective tax rate (hereafter, the tax change component of earnings) are persistent for future earnings.[1] In addition, this study investigates to what extent investors incorporate the forecasting implications of the tax change component of earnings into stock prices. The tax change component of earnings is likely to have implications for earnings persistence for a number of reasons. First, income taxes are of such magnitude that small changes in the effective tax rate (ETR) can lead to substantial changes in earnings. Second, corporate managers are placing greater emphasis on ETRs as in-house tax departments are increasingly viewed as profit centers (U.S. Department of the Treasury 1999; McGill and Outslay 2004).[2] Finally, recent academic literature, as well as the financial press, has identified ETR manipulation as an earnings management tool. Dhaliwal et al. (2004) and Comprix et al. (2006) find that firms lower ETRs when earnings absent tax management fall short of analyst forecasts. Myers et al. (2005) indicate that managers use ETRs to smooth reported EPS.

The tax change component of earnings captures the aggregate effect of ETR reconciliation items, including permanent book-tax differences. However, ETR reconciliation items increasingly reflect the effects of strategic tax-planning/tax-optimization activities, which are designed to provide sustainable long-term benefits to earnings. For example, management of Tyco urged shareholders to reject a proposal to move its parent company back to the U.S. from Bermuda, stating that the impact on the company's current and future earnings (from an estimated seven-percentage-point increase in its current and future ETR) could have caused a $4 to $5 billion decrease in market capitalization (McGill and Outslay 2004). Observers speculate that items like these are a major reason for the increase in book-tax differences in the 1990s (Reason 2002).

The collapse of Enron, the rise of global tax optimization (planning), and the renewed use of corporate tax shelters has resulted in increased scrutiny of book-tax differences and their implications for earnings quality and financial performance. My study contributes to the growing literature that examines the persistence and pricing implications of book-tax differences by directly examining how taxes contribute to earnings prediction, which is consistent with the view that one task of accounting research is to discover what information in the financial statements projects future earnings (Penman 1992).[3] Understanding the differential persistence of the tax change component and its forecasting implications should aid market participants in profitability analysis and in income forecasting.

---

[1] A subtle difference between this study and studies that examine taxes as one of several fundamental signals used in financial statement analysis (Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998; Amir et al.2000) is how the effective tax rate (ETR) and the tax change component is defined. I define the ETR as total tax expense divided by pretax income and the tax change component as pretax income multiplied by the change in the ETR. Lev and Thiagarajan (1993) use current tax expense only in the numerator of the ETR and Abarbanell and Bushee (1997, 1998) and Amir et al. (2000) use pretax income adjusted for amortization of intangibles in the denominator of the ETR. Additionally, Abarbanell and Bushee (1997, 1998) and Amir et al. (2000) interact the change in ETR with the change in after-tax EPS in order to determine the earnings effect of ETR changes, which differs from my definition of the tax change component.

[2] Beginning in 1995, Enron's tax division used a series of structured transactions to increase company profits. The deals were intended to create $1.9 billion in earnings over 30 years (Witt and Behr 2002).

[3] This approach differs from that of Guenther and Jones (2003), who study the relationship between contemporaneous stock returns and annual ETR-related earnings changes in order to determine whether annual ETR-related earnings changes are value-relevant. An important assumption in Guenther and Jones (2003) is that the value-relevance of ETR-related earnings is indicative of future earnings or growth opportunities of the firm. Guenther and Jones (2003) do not test whether ETR-related earnings changes are associated with realized future earnings, which is the primary focus of this paper.

Prior literature either ignores or aggregates taxes with other earnings components (Sloan 1996; Hribar 2001; Nissim and Penman 2001; Barth et al. 2004) or examines the implications of information contained in temporary book-tax differences (deferred tax expense) for the quality of pretax earnings (Hanlon 2005). An additional branch of research examines the informativeness of estimated taxable income to capital markets (Shevlin 2002; Hanlon et al. 2005; Lev and Nissim 2004).

In contrast to these studies, my decomposition of earnings treats taxes like any other earnings component (i.e., cash flows, changes in accounts receivable, changes in inventory, etc.), which allows me to directly test a maintained hypothesis that earnings generated by ETR changes are transitory or attempts at aggressive accounting by management.

I find that there is a positive, significant association between the tax change component of earnings and future earnings, which disputes the broad notion advanced by prior literature (Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998; Amir et al. 2000) that ETR-related earnings changes are completely transitory. In contrast to these earlier studies, as well as concurrent research by Guenther and Jones (2003), I use the interim reporting requirements of APB No. 28 (APB 1973) and FASB Interpretation No. 18 (FASB 1977) to further decompose the tax change component into an initial and a revised portion based on the first quarter estimate of the annual ETR. I find that the initial tax change component is more persistent for future earnings than the revised tax change component.

This study also complements concurrent research that suggests tax-related earnings information is priced by the market (Hanlon 2005; Gleason and Mills 2004; Guenther and Jones 2003; Lev and Nissim 2004) by directly testing whether investors appropriately incorporate the implications of ETR-related earnings changes for future earnings into market prices.

Results from market tests indicate that investors underweight the persistence of the tax change component and the mispricing appears to be driven by the transitory nature of the revised tax change component. Supplemental tests confirm that the results obtained from the market tests are likely a result of mispricing rather than alternative explanations such as omitted risk factors or previously documented anomalies.

I organize the remainder of the paper as follows. Section II presents an overview of corporate income tax disclosures, emphasizing the issues surrounding the ETR reconciliation and how current reporting requirements affect my empirical proxies. Section III develops hypotheses. Section IV outlines the research design; Section V presents results and supplemental analyses. Section IV concludes.

II. CORPORATE INCOME TAX DISCLOSURES

Effective Tax Rates

Disclosure

Recent accounting standards provide guidance regarding ETR disclosures. Statement of Financial Accounting Standards (SFAS) No. 96 (FASB 1987, ¶28) and No. 109 (FASB 1992, ¶47), *Accounting for Income Taxes*, require companies to disclose a reconciliation of the amount of income tax expense that would result from applying the domestic federal statutory rate to pretax income from continuing operations to the reported amount of income tax expense attributable to continuing operations for the year. When a firm reconciles its statutory rate to the ETR, it must report any item that results in a material difference between the statutory rate and the ETR.[4]

Reconciliation Adjustments

Permanent differences cause the ETR to deviate from the statutory rate. Permanent differences are caused by items that enter into pretax book income but never enter into taxable income and items that enter into taxable income but never into pretax book income.[5] Common examples of permanent differences include:

- *Items recognized in pretax book income but not in taxable income*: Interest on tax-exempt municipal bonds and life insurance proceeds on an officer of the corporation, qualified stock options accounted for using the fair value method, penalties and fines, key employee life insurance premiums, and most goodwill and in-process research and development (IPRD) in acquisitions.[6]
- *Items recognized in taxable income but not in pretax book income*: Percentage depletion in excess of cost depletion and the dividends-received deduction.

Additionally, certain items that are technically temporary differences have the same effect as a permanent difference because a deferred tax liability is not recognized for these items (SFAS No. 109 ¶31a–d). One example is the undistributed earnings of a foreign subsidiary consolidated for book purposes that management intends to reinvest indefinitely.[7]

Tax rate adjustments also cause the ETR to deviate from the statutory rate. Tax rate adjustments are special tax rates or tax credits that affect the statutory rate. Tax rate adjustments are not permanent or temporary differences as they do not affect pretax income or taxable income (Weber and Wheeler 1992) and include:

- Tax credits (e.g., the research and development credit or the foreign tax credit).
- Different tax rates on foreign income.
- Tax expense that includes both state and local income taxes, net of federal tax benefit.

Finally, the ETR reconciliation includes adjustments not only for the tax effects of pretax income from continuing operations, but also for the tax effects of other items attributable to continuing operations, including:

---

[4] If no individual reconciling item amounts to more than 5 percent of the amount computed by multiplying the income before tax by the applicable statutory Federal income tax rate, and the total difference to be reconciled is less than 5 percent of such computed amount, then no reconciliation need be provided unless it would be significant in appraising the trend of earnings. Reconciling items that are individually less than 5 percent of income tax expense may be aggregated in the reconciliation (SEC Regulation S-X §210.4-08(h)(2)).

[5] Permanent differences are defined in ¶33 of APB No. 11. However, the term "permanent difference" is not used in SFAS No. 109. Instead, SFAS No. 109 (¶14) states that certain differences between the tax basis and accounting basis of assets and liabilities will not result in taxable or deductible amounts in future years, therefore no deferred tax liability or asset should be recognized.

[6] Until 1993, goodwill was always a permanent difference; it was amortized for book purposes for a period not to exceed 40 years and was not deductible on the tax return. After 1993, goodwill may be amortized over 15 years for tax purposes (IRC section 197) as long as the tax bases of the acquired firm's assets are stepped-up. A step-up occurs almost exclusively in subsidiary sales or acquisitions of conduits. In these cases, goodwill results in a temporary difference. Goodwill created through the acquisition of a freestanding C-corporation, where basis step-ups are rare, usually results in a permanent difference (Scholes et al. 2004). IPRD is written off immediately for financial reporting purposes. Similar to goodwill, IPRD is not tax deductible as long as no tax basis exists.

[7] Undistributed earnings of subsidiaries consolidated for financial reporting purposes are recognized in pretax book income currently whether such subsidiaries are foreign or domestic. These earnings would usually be recognized in taxable income when distributed, resulting in a temporary difference. However, if a foreign subsidiary intends to reinvest its undistributed earnings indefinitely, no deferred tax liability is recorded, which effectively treats the undistributed earnings as a permanent difference.

- deferred tax asset valuation allowance account adjustments;
- changes in tax laws or tax rates;
- changes in the tax status of an enterprise (e.g., from a partnership to a corporation); and
- tax loss carryforward benefits.

The tax change component of earnings reflects the aggregate effect of ETR reconciling items; therefore, the extent to which the tax change component of earnings persists and aids in forecasting future earnings provides an indication of how ETR reconciliation items can be used in financial statement analysis.

Interim Reporting Requirements

GAAP mandates that the quarterly reporting period be viewed as an integral part of the annual reporting period.[8] APB Opinion No. 28 (APB 1973), SFAS No. 3 (FASB 1974), and FASB Interpretation No. 18 (FASB 1977) require firms to estimate many annual operating expenses and then allocate these estimates to interim periods based on forecasted figures. As the fiscal year progresses, firms revise estimates and record estimation errors from earlier quarters (Rangan and Sloan 1998).

Income taxes are one type of expense that is subject to the interim reporting requirements. Firms compute quarterly tax expense by estimating the ETR that is expected to be applicable for the full fiscal year. The estimated annual ETR is then applied to the year-to-date ordinary income at the end of each interim period, defined as pretax income before extraordinary items, discontinued operations, and cumulative effects of changes in accounting principles (FASB Interpretation No. 18, 1977, ¶9), to compute the year-to-date tax expense. The year-to-date tax expense of the prior quarter is then subtracted from the year-to-date tax expense of the current quarter to arrive at the current quarter's tax expense. This procedure is illustrated in the following footnote from Hewlett-Packard's Q3 2000 10-Q:

> HP's effective tax rate was 23% for the first nine months of fiscal 2000, reflecting the tax rate expected for the full fiscal year. The tax rate was 21.3% in the third quarter of 2000 due to a reduction in the annual effective tax rate from 24% to 23%. In 1999, the full-year effective tax rate was 26%, and the rate was 26.5% in both the third quarter and first nine months of the year. The year-to-year decrease in HP's effective tax rate was primarily the result of changes in the mix of our pretax earnings in various tax jurisdictions throughout the world.

I use the interim reporting requirements to decompose the tax change component into an initial and revised portion based on the first quarter estimate of the annual ETR. This decomposition allows me to develop and test hypotheses regarding the differential persistence and forecasting implications of the tax change component of earnings.

III. RESEARCH QUESTIONS AND HYPOTHESES

The Persistence and Forecasting Implications of the Tax Change Component of Earnings

The first objective of my paper is to document whether the tax change component is persistent for future earnings. Variations of the tax change component have been studied previously as part of a set of fundamental signals used by financial analysts to predict future performance (Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998). The underlying hypothesis in these papers is that analysts view earnings generated from ETR changes as transitory.[9] Abarbanell and Bushee (1997) find that ETR-related earnings changes are less persistent for future earnings changes and long-term earnings growth than aggregate earnings changes from 1983–1990, but do not directly test (or report) the extent to which ETR-related earnings changes are associated with future earnings changes. Despite this, Abarbanell and Bushee (1997) state that the ETR variable captures more than transitory effects, possibly unidentified risk factors or structural changes.

---

[8] The integral approach was selected over a discrete approach for interim reporting. Under the discrete approach, each interim quarter is viewed as a separate accounting period, therefore interim accruals and deferrals should reflect the same principles employed for annual reports. The discrete approach applies the same expense recognition principles to both interim and annual reports with no special interim accruals or deferrals allowed.

The view that earnings generated from ETR changes contain some transitory elements is reasonable because the ETR can be used for period-specific earnings management (i.e., beating analyst forecasts, avoiding a loss, or avoiding a decrease in earnings). Recent research examines earnings management through aggregate ETR changes (Comprix et al. 2006; Dhaliwal et al. 2004; Myers et al. 2005) and specific ETR reconciling items such as the DTVA (Miller and Skinner 1998; Visvanathan 1998; Bauman et al. 2001; Burgstahler et al. 2003; Frank and Rego 2003; Schrand and Wong 2003), permanently reinvested foreign earnings (Krull 2004), intangible assets (Nelson et al. 2003), and the tax cushion (Smith 2001; Gleason and Mills 2002; Nelson et al. 2003). The managed portion of earnings could be transitory (or less persistent), resulting in a lower coefficient on managed earnings in predicting future earnings, compared to the coefficient on unmanaged earnings (Xie 2001; Joos et al. 2003; Hanlon 2005).

However, some of the most prominent ETR reconciliation items, including foreign tax rate differentials and state taxes, likely reflect long-term (and therefore persistent) strategic tax-planning activities as well as contain forward-looking information that may shed light on the future prospects of the firm.[10] Mills et al. (1998) show that there is a positive relationship between investments in tax planning and long-term strategy indicators such as the presence of foreign assets, number of entities, and capital intensity. Bodnar and Weintrop (1997) find that foreign operations provide greater growth opportunities than domestic operations and Thomas (2000) finds that foreign earnings changes are more useful than domestic earnings changes in predicting future earnings changes. If greater growth opportunities lead to more persistent earnings (Collins and Kothari 1989), ETR-related earnings in foreign jurisdictions may be more permanent depending on the extent to which the company maintains or expands its operations in foreign jurisdictions.

Further, state tax-planning activities have grown tremendously in the 1990s, in part due to fewer (and more costly) federal tax-planning opportunities, economic-development-oriented tax incentive competition among states, and more aggressive state tax compliance due to state fiscal problems. Gupta and Mills (2002) show that state ETRs decrease (at a decreasing rate) and then increase (at an increasing rate) as a function of the number of state tax returns filed by the firm. This information would be reflected in the ETR and would increase the persistence of the tax change component of earnings through the extent to which the benefits of state tax planning result in permanent tax savings.

---

[9] This hypothesis, which originated in Lev and Thiagarajan (1993), likely comes from the following sources: (1) various accounts in O'Glove (1987, 72) where a decrease in the ETR is characterized as a nonrecurring/nonoperating item, (2) a survey by Siegel (1982), which indicates that analysts, accountants, and financial managers think that an increase in earnings from the utilization of a net operating loss carryforward represents a deterioration of earnings quality, and (3) articles in publications such as the *Wall Street Journal* that discussed characteristics of company earnings releases.

[10] Strategic tax-planning activities also include ETR "optimization" strategies such as transferring intangibles offshore, transfer pricing, structuring low-tax contract manufacturing operations, using holding companies to minimize withholding taxes on dividends, and engaging in intercompany financing techniques like debt push-downs.

The persistence of the tax change component could also be affected by the renewed use of tax shelters starting in the mid-1990s.[11] While most tax shelters are thought to provide one-time benefits, evidence suggests that some tax shelters could have more persistent effects on earnings. For example, Graham and Tucker (2005) find that the average tax shelter used by firms in their sample was active for approximately five years. While most tax-sheltering activity is likely not as persistent as long-term strategic tax planning, there could be a positive effect on the persistence of the tax change component from tax sheltering to the extent that certain tax shelters provide benefits over multiple years.

Since the tax change component of earnings (as does any earnings component) has effects that are expected to be either persistent or transitory, whether the tax change component of earnings persists or aids in forecasting future earnings depends on the relative proportion of persistent and transitory effects. The previous discussion suggests that the tax change component of earnings primarily reflects a combination of long-term strategic tax planning (expected to be relatively persistent), tax-sheltering activities, and period-specific earnings management activities (both expected to be relatively transitory). This raises the following question:

> **RQ$_1$:** Is the tax change component of earnings persistent for future earnings (i.e., useful in forecasting future earnings incremental to earnings excluding the tax change)?

The Persistence and Forecasting Implications of the Tax Change Component of Earnings in Interim Periods

I next consider the persistence of the tax change component of earnings in interim periods. The results of prior research on the time-series properties of quarterly earnings are consistent with the notion that interim earnings become less persistent (more transitory) as the fiscal year progresses.[12] Thus, even without considering aspects specific to tax expense, one might expect the tax change component early in the year to be more persistent than any revisions.

However, the interim reporting requirements for income tax expense suggest additional arguments for the tax change component to be more or less persistent. APB No. 28 requires managers to make their best estimate of the firm's ETR that is expected to be applicable for the full fiscal year when computing interim income tax expense. The estimated annual ETR should reflect anticipated tax credits, foreign tax rates, percentage depletion, capital gains rates, and other available tax-planning alternatives (FASB Interpretation No. 18, 1977, ¶8).

---

[11] Tax advisors and corporate directors regard a transaction with low visibility that results in a permanent difference as the perfect tax shelter (U.S. Department of Treasury 1999; Hanlon 2003; Boynton and Mills 2004; McGill and Outslay 2004). These shelters would affect the ETR and, thus, the tax change component of earnings.

[12] Collins et al. (1984) find that the forecast errors tend to increase as the fiscal year progresses, while Bathke and Lorek (1984) find that time-series forecast errors are larger in the fourth quarter. Salamon and Stober (1994) and Kross and Schroeder (1990) present evidence that earnings response coefficients are significantly smaller for fourth quarter earnings than for interim earnings, which is consistent in part with the existence of fiscal year-end discretionary accruals from earnings management. Lipe and Bernard (2000) find that the differential stock-price response to fourth quarter earnings is most likely due to increased recognition of transitory earnings in the fourth quarter, opposed to the traditional explanation of the annual audit correcting interim reporting errors. The Lipe and Bernard (2000) finding is consistent with Burgstahler et al. (2002), who find that special items, which proxy for transitory earnings, occur more frequently during quarters later in the fiscal year, with about 41 percent occurring in the fourth quarter, and Livnat (2003), who finds that fourth quarter earnings have lower persistence than earnings from the first through third quarters and the decline in persistence is attributed mostly to extremely negative earnings surprises.

Interviews with current and former tax partners suggest that it is likely that the items that have the most material (and persistent) effect on management's estimate of the annual ETR are anticipated and incorporated into the initial (Q1) annual ETR estimate.[13] For example, it may be that early in the year, managers make assumptions regarding the relative long-term strategic mix of income from various sources (i.e., domestic versus foreign earnings, taxable versus tax-advantaged income) and use these estimates (in part) to arrive at the initial annual estimate of the ETR. As period-specific information becomes available in subsequent quarters (Q2, Q3, Q4), managers may revise the annual ETR estimate for period-specific deviation in the income weights.[14,15]

Anecdotal evidence suggests that analysts often view revisions in estimated annual ETRs as attempts at aggressive accounting, and prior research finds that firms change ETRs to beat analysts' forecasts or to smooth earnings, consistent with earnings management (Dhaliwal et al. 2004; Comprix et al. 2006; Myers et al. 2005). However, it is ultimately an empirical question as to whether the initial (Q1) tax change component of earnings behaves differently than subsequent revisions (Q2, Q3, Q4) in the tax change component of earnings.[16] If tax-planning (earnings management) activities are more persistent (transitory) and if the initial tax change component reflects more tax-planning activities than earnings management activities, then the initial tax change component will be more persistent (or less transitory). Likewise, if the revised tax change component reflects more earnings management activities, then the revised tax change component will be less persistent. Stated formally:

H_1: The initial (Q1) tax change component of earnings is more persistent for future earnings than subsequent revisions (Q2, Q3, Q4) in the tax change component of earnings.

---

[13] The partners stressed that managers are aware of the major permanent differences and other ETR reconciliation items when firms predict end-of-year book income. Furthermore, managers do not like surprises from the tax department, which often leads to conservative initial ETR estimates. When an ETR is changed after the initial estimate is made at the end of the first quarter (Q1), it is usually due to one of the following reasons: (1) there is a surprise, such as an audit settlement that is better than what was previously anticipated; (2) new tax planning; (3) there is a return to reality (i.e., the initial conservative ETR estimates from Q1 are "trued up" as future earnings are realized); (4) earnings management; or (5) significant variations in pretax earnings (especially pretax earnings close to zero). Of these reasons, the partners agreed that new tax planning is least likely to occur.

[14] Additional insight into this phenomenon (albeit with different income sources) can be gained from the Q3 2000 MD&A of Viad (2000): "The relatively low ETR compared to the statutory federal rate is primarily attributable to tax-exempt income from Viad's Payment Services businesses. APB Opinion No. 28 requires that income taxes be provided based on the estimated ETR expected to be applicable for the entire fiscal year, with an adjustment of the annual ETR made each quarter. During the third quarter, Viad determined that the estimated annual ETR for 2000 is expected to be lower than in prior periods due to higher than previously expected tax-exempt income in proportion to total pretax income, resulting from rapid growth in investments in tax-exempt securities in the Payment Services segment along with lower operating income in the Convention and Event Services Segment. Accordingly, the adjustment of the ETR was made in the third quarter of 2000."

[15] This could also be an additional way the ETR is managed. In fact, when companies make interim ETR revisions, one of the most common explanations for the revision is a change in the estimated mix of income from foreign and domestic sources (for example, see the Q3 2000 10-Q of Hewlett-Packard, a part of which is presented in Section II).

[16] Unless a tax shelter provides benefits for multiple years, it is unlikely that management's estimate of the annual ETR will reflect the tax benefits from the shelter; therefore, it is not likely that this type of activity will affect the persistence of the initial and revised tax change component in a different manner.

Valuation Implications of the Tax Change Component of Earnings

My final objective is to determine if market participants recognize the forecasting implications of the tax change component. Until recently, the empirical evidence (Lipe 1986; Lev and Thiagarajan 1993; Abarbanell and Bushee 1997, 1998) on the value-relevance of ETR-related earnings changes has been inconclusive and inconsistent. Guenther and Jones (2003) reconcile these studies and determine that the conflicting results are due primarily to how the tax change component is measured in each study (see footnote 1).

Guenther and Jones (2003) document a significant, positive association between contemporaneous abnormal returns and the tax change component and show that this association varies cross-sectionally and over time. Guenther and Jones (2003) hypothesize that the variation in the magnitude of the association between contemporaneous returns and the tax change component is due to the persistence of the tax change component in different contexts.[17] However, this expectation assumes that the market is efficient with respect to the implications of the tax change component for future earnings.[18]

Despite the recently documented value-relevance of the tax change component, there is conflicting evidence about the efficient use of this information by market participants. For example, Abarbanell and Bushee (1997) find that analysts overreact to annual ETR-related earnings changes before the implementation of SFAS No. 109, while Chen and Schoderbek (2000) find that analysts failed to incorporate the deferred tax adjustment required by SFAS No. 109 into their earnings forecasts when corporate tax rates increased in 1993. Bauman and Shaw (2005) find that analysts' quarterly earnings forecasts incorporate (underreact to) the earnings effects of ETR increases and small ETR decreases (the earnings effects of large ETR decreases) in previous quarters after the implementation of SFAS No. 109.

Evidence of analyst forecast inefficiency raises the question of whether stock prices fully reflect the persistence of the tax change component of earnings. Further complicating matters, research indicates that investors inefficiently utilize information contained in analysts' earnings forecasts. Mendenhall (1991), Walther (1997), and Elgers et al. (2001, 2003) document that investors underweight information contained in analysts' forecasts, while LaPorta (1996) and Bradshaw et al. (2001) find that investors overweight information contained in analysts' forecasts. Therefore, it is not entirely clear that investors' expectations of the forecasting implications of the tax change component will be consistent with the implications of the tax change component for future earnings. This leads to the final hypotheses:

H_{2a}: The earnings expectations embedded in stock prices fail to fully reflect the persistence of the tax change component of earnings.

---

[17] The positive coefficient on the tax change component is significantly smaller after firms' adoption of SFAS No. 109 in 1993. Further, in the post-SFAS No. 109 period the coefficient on the tax change component is significantly smaller for firms reporting the utilization of a net operating loss and for firms recording restructuring charges or asset impairments, and is significantly larger for firms engaged in acquisitions and for firms that experience a change in the percentage of foreign earnings.

[18] The results of Lev and Nissim (2004) and Hanlon (2005) weaken this assumption. Lev and Nissim (2004) document that the market does not fully reflect the information contained in a tax fundamental that captures aggregate book-tax differences. Hanlon (2005) finds that investors underestimate the persistence of the cash flow component of earnings for firm-years with large positive temporary book-tax differences (i.e., book income greater than taxable income) and overestimate the persistence of accruals when firm-years contain large negative temporary book-tax differences (i.e., book income less than taxable income).

H_{2b}: The earnings expectations embedded in stock prices fail to fully reflect the persistence of the initial (Q1) and revised (Q2, Q3, Q4) tax change component of earnings.

IV. RESEARCH DESIGN

Model Development

To focus on the effect of an ETR change on earnings, I follow Lev and Thiagarajan (1993) and decompose the annual change in earnings, $\Delta E_t = E_t - E_{t-1}$, into two components:

(1) The change in pretax earnings (ΔPTE_t), using last year's effective tax rate (ETR_{t-1}): $\Delta PTE_t(1 - ETR_{t-1})$, and

(2) The effect of the annual ETR change on current pretax earnings: $PTE_t(ETR_{t-1} - ETR_t)$:

$$\Delta E_t = \Delta PTE_t(1 - ETR_{t-1}) + PTE_t(ETR_{t-1} - ETR_t). \tag{1}$$

In Equation (1), $PTE_t(ETR_{t-1} - ETR_t)$, which represents earnings generated by a change in the annual ETR, is my measure of the tax change component of earnings.[19] *ETR* is a firm's effective tax rate, measured as total income tax expense divided by *PTE*. Interim (annual) Compustat data items used for each variable are: Earnings = 8 (18), Income Tax Expense = 6 (16), and Pretax Earnings = 23 (170). Following Gupta and Newberry (1997), the *ETR* is set to 1 when the ETR is greater than 100 percent, and to 0 when the ETR is negative.

The Persistence and Forecasting Implications of the Tax Change Component of Earnings

I examine the persistence of the tax change component of earnings for future earnings using the following equation (firm subscripts are omitted):

$$E_{t+1} = \gamma_0 + \gamma_1 ATE_t + \gamma_2 TCC_t + \nu_{t+1} \tag{2}$$

where *E* represents income before extraordinary items, *ATE* represents aggregate earnings excluding the tax change component [$ATE = PTE_t(1 - ETR_{t-1})$], and *TCC* represents the tax change component of earnings [$TCC = PTE_t(ETR_{t-1} - ETR_t)$].[20] I scale all earnings component variables by contemporaneous average total assets (Compustat annual data item

---

[19] The following example, adapted from Lev and Thiagarajan (1993), should provide intuition for this measure. Assume:

	t	$t-1$
Pre-tax income	120	100
ETR	0.25	0.40
Net income	90	60

Following Equation (1), the net income change, 30, consists of the pretax earnings change net of the prior tax rate, 20 × (1 − 0.40) = 12, and the effect of the tax rate change, 120 (0.40 − 0.25) = 18. I refer to the latter component of the earnings change as the tax change component.

[20] My partitioning of earnings into a tax component (*TCC*) and a pretax component (*ATE*) generally assumes that the pretax component is unrelated to the tax component. I acknowledge that some tax-planning choices directly affect pretax earnings through implicit taxes. Although it is possible that some of the pretax persistence is related to implicit taxes, I assume that the magnitude of any implicit tax effects is too small to affect my inferences.

#6) to control for differences in size across firms and over time and to allow for cross-sectional comparability.

The Persistence and Forecasting Implications of the Tax Change Component of Earnings in Interim Periods

The interim reporting requirements of APB Opinion No. 28 and FASB Interpretation No. 18 allow me to further decompose the tax change component of earnings into an initial and a revised component based management's first quarter estimate of the ETR that is expected to be applicable for the full fiscal year. Hypothesis 1 predicts that Equation (2) is misspecified since it implicitly constrains the coefficients on the initial and revised tax change components of earnings to be equal. The specification implied by H_1 is:

$$E_{t+1} = \delta_0 + \delta_1 ATE_t + \delta_2 INTCC_t + \delta_3 REVTCC_t + \nu_{t+1} \tag{3}$$

where *INTCC* represents the initial tax change component of earnings $[PTE_t(ETR_{t-1} - ETRQ1_t)]$ and *REVTCC* represents the revised tax change component of earnings $[PTE_t(ETRQ1_t - ETR_t)]$. $ETRQ1_t$ is the first quarter estimate of the annual ETR. Hypothesis 1 predicts $\delta_2 > \delta_3$. The larger coefficient predicted on the initial relative to the revised tax change component of earnings reflects the higher persistence expected from the initial tax change component of earnings. To mitigate concern that any differential persistence is driven by asset growth, I scale the dependent variables in Equations (2) and (3) by lagged average total assets so the scalar is the same as that used for the independent variables (Fairfield et al. 2003).

Finding a significant positive relation for any coefficient in Equations (2) or (3) indicates that the component is incrementally informative in predicting future earnings. For each earnings component, a slope coefficient of 0 indicates that the earnings component is purely transitory, while a coefficient of 1 indicates that the earnings component follows a random walk. A slope coefficient greater than 1 indicates growth in the component.

Valuation Implications of the Tax Change Component of Earnings

My final objective is to test whether the market fully incorporates the forecasting implications of the initial, revised, and annual tax change component of earnings. To test H_{2a} and H_{2b}, I estimate the following systems of equations in accordance with the Mishkin (1983) framework:

$$E_{t+1} = \gamma_0 + \gamma_1 ATE_t + \gamma_2 TCC_t + \nu_{t+1} \tag{2}$$

$$AR_{t+1} = \phi_0 + \phi_1(E_{t+1} - \gamma_0 - \gamma_1^* ATE_t - \gamma_2^* TCC_t) + \varepsilon_{t+1} \tag{4}$$

$$E_{t+1} = \delta_0 + \delta_1 ATE_t + \delta_2 INTCC_t + \delta_3 REVTCC_t + \nu_{t+1} \tag{3}$$

$$AR_{t+1} = \omega_0 + \omega_1(E_{t+1} - \delta_0 - \delta_1^* ATE_t - \delta_2^* INTCC_t - \delta_3^* REVTCC_t) + \varepsilon_{t+1} \tag{5}$$

where AR_{t+1} = the abnormal return in year $t+1$, measured as the 12-month buy-and-hold return beginning four months after the fiscal year-end minus the buy-and-hold value-weighted return of a benchmark portfolio over the same 12-month period (Daniel et al. 1997). I create 125 (5 × 5 × 5) size/book-to-market/prior year return portfolios each year to control for a set of common risk factors: size, book-to-market, and momentum.[21] I measure each characteristic at the end of the most recent fiscal year prior to portfolio formation.

Equation (2) is a *forecasting* equation that estimates the forecasting coefficients (γ) of aggregate earnings excluding the tax change and the tax change component for predicting future earnings.[22] Each coefficient is a measure of the persistence of an earnings component for future earnings. Equation (4) is a *valuation* equation that estimates valuation coefficients (γ^*) that the market assigns to earnings components. Mishkin (1983) suggests that Equation (4) provides an estimate of the market's *perceived* persistence of earnings and the tax change component. Because earnings and the tax change component of earnings are public information, market efficiency requires $\gamma_1 = \gamma_1^*$, and $\gamma_2 = \gamma_2^*$. If any pair of coefficients is not equal, then the market's perception of earnings persistence differs from the actual earnings persistence implied by the specification of the forecasting equation.

I estimate the system of Equations (2) and (4) jointly in two stages using iterative generalized nonlinear least squares (Mishkin 1983). In the first stage, I jointly estimate an unconstrained system of Equations (2) and (4). In the second stage, I impose the rational pricing constraints $\gamma_q = \gamma_q^*$ ($q = 1$ or 2) and re-estimate Equations (2) and (4) jointly. I test market efficiency with the likelihood ratio statistic described in Mishkin (1983). The specific likelihood ratio test is $2n\log(SSR^c/SSR^u) \sim \chi^2(q)$ where n = the number of observations, q = the number of rational pricing constraints, SSR^c = the sum of squared residuals from the constrained weighted system, and SSR^u = the sum of squared residuals from the unconstrained weighted system.

Sample Selection Procedures and Sample Profile

My sample consists of all U.S. Compustat firm-years between 1994 and 2001 with nonmissing values of annual and quarterly assets, total income tax expense, and pretax income (36,442 observations). Due to the difficulty in interpreting ETRs when a firm has negative pretax income, I exclude 13,581 firm-years with negative pretax income. Then, I delete 1,591 mutual funds, trusts, REITs, limited partnerships, and other flow-through entities due to different reporting requirements and varying earnings management and tax-planning incentives. Next, I exclude 7,753 firm-years with missing CRSP return data. Finally, to remove the effects of outliers from the data, I drop 822 firm-years with the highest and lowest 1 percent of the values for each variable in year t (Kothari and Zimmerman 1995; Collins et al. 1997; Fama and French 1998; Barth et al. 1999).[23] My final sample consists of 12,695 firm-years, which are distributed uniformly throughout the sample period.

V. EMPIRICAL RESULTS

Descriptive Statistics

Table 1 reports descriptive statistics for the variables used in this study. Panel A describes the variables that I use in the calculation of the tax change component of earnings. The average annual ETR in the prior and current year is 35.68 percent and 35.47 percent, respectively, while the average ETR in Q1, which represents management's expectation of the annual ETR, is 35.93 percent. $ETRQ1_t$ is significantly higher than ETR_{t-1} and ETR_t

---

[21] Firms must have calendar year-end market capitalization values available on CRSP and positive fiscal year-end book value data (annual data item #60) available on Compustat.

[22] The discussion of the Mishkin (1983) procedure for Equations (2) and (4) generalizes to Equations (3) and (5).

[23] Extreme returns or earnings in year $t+1$ are not eliminated since this would introduce hindsight bias in the results (Thomas 2000). I also run all of my tests on data that has been winsorized rather than trimmed. The inferences from all of my tests are unchanged when I use the winsorized data.

TABLE 1
Descriptive Statistics

Variable	Mean	Standard Deviation	1st Quartile	Median	3rd Quartile
Panel A: Financial Variables[a]					
$Assets_t$	6306.79	34104.66	163.003	553.621	2161.740
PTE_t	275.72	1002.81	11.679	40.740	147.345
ETR_t	0.3568	0.0840	0.3304	0.3676	0.3939
ETR_{t-1}	0.3547	0.0788	0.3318	0.3690	0.3945
$ETRQ1_t$	0.3593	0.0756	0.3389	0.3716	0.3960
Panel B: Regression Variables[b]					
E_{t+1}	0.0659	0.0544	0.0243	0.0534	0.0932
AR_{t+1}	0.0409	0.6180	−0.2951	−0.0393	0.2639
TCC_{t+1}	−0.0004	0.0098	−0.0014	0.0000	0.0015
ATE_t	0.0700	0.0532	0.0247	0.0534	0.0938
TCC_t	−0.0002	0.0044	−0.0011	0.0000	0.0010
$INTCC_t$	−0.0005	0.0042	−0.0012	−0.0001	0.0002
$REVTCC_t$	0.0003	0.0029	−0.0003	0.0001	0.0015
Panel C: Risk and Anomaly Control Variables[c]					
MVE_t	3373.70	15896.93	104.35	397.86	1506.92
BMR_t	0.2701	1.8811	−0.1415	0.1567	0.6970
$PBHR_t$	0.1807	0.5498	−0.1228	0.1071	0.3701
$BETA_t$	0.8789	0.5997	0.4732	0.8089	1.1958
$SPEC_t$	−0.0024	0.0166	0.0000	0.0000	0.0000
$DTAX_t$	−0.0018	0.0115	−0.0065	−0.0009	0.0033
VOL_t	95.1764	113.4364	33.7918	61.2824	108.9021
EP_t	0.0698	0.0398	0.0457	0.0642	0.0858
$PTXACC_t$	−0.0244	0.0761	−0.0642	−0.0306	0.0057
$PTXCF_t$	0.1462	0.1025	0.0840	0.1377	0.2024

Panel D: Industry Composition[d]

Industry Membership	Mean *TCC*	Mean *INTCC*	Mean *REVTCC*
Mining and Construction	−3.1851	−6.5646	3.3795
Food	1.1964	−2.0408	3.2372
Textiles, Printing, Publishing	−0.0920	−3.4453	3.3534
Chemicals	0.0651	−7.4574	7.5225
Pharmaceuticals	−0.4225	−8.2941	7.8716
Extractive Industries	−15.0310	−14.3429	−0.6881
Durable Manufacturers	−2.0190	−8.1600	6.1410
Computers	−4.6170	−7.4794	2.8624
Transportation	−2.3980	−1.1700	−1.2280
Utilities	−3.4671	−5.9137	2.4466

(*continued on next page*)

TABLE 1 (Continued)

Industry Membership	**Mean** *TCC*	**Mean** *INTCC*	**Mean** *REVTCC*
Retail	−1.3513	−4.1479	2.7967
Financial Institutions	−1.6000	−1.9117	0.3117
Insurance and Real Estate	−4.6166	−7.9737	3.3571
Personal Services	−3.8471	−6.2612	2.4141
Professional Services	−3.7826	−5.6880	1.9054
Other	−2.1882	−7.6665	5.4783
Total	−2.9597	−6.1573	3.1976

[a] Panel A variables are defined as follows: *Assets* = total assets (Compustat annual item #6); *PTE* = Pretax income (Compustat annual item #170); *ETR* = Effective tax rate (Compustat annual item #16 ÷ *PTE*); and *ETRQ1* = Q1 Effective tax rate (Compustat quarterly item #16 + Compustat quarterly item #23).

[b] Panel B variables are defined as follows: *E* = income before extraordinary items (Compustat annual item #18). *AR* = the 12-month buy-and-hold security return beginning four months after the fiscal year end minus the value-weighted return of the benchmark size/book-to-market/prior year return portfolio over the same 12-month period. The benchmark portfolios are created by first sorting stocks into quintiles based on their capitalization values (*MVE*) at the beginning of the year of portfolio formation. NYSE breakpoints are used so that there are an equal number of NYSE stocks in each quintile. Next, within each of the size quintiles, stocks are sorted into quintiles based on their industry-adjusted book-to-market ratios. The book-to-market (*BMR*) ratio is defined as the book value at the end of the fiscal year prior to the year of portfolio formation divided by capitalization value at the beginning of the calendar year of portfolio formation. Book-to-market ratios are industry adjusted by subtracting the mean industry book-to-market ratio over the sample period from the individual stock's book-to-market ratio, where industries are defined along two-digit SIC codes. Finally, within the 25 size/book-to-market portfolios, stocks are sorted based on their prior 12-month return (*PBHR*) ending one month prior to portfolio formation. *TCC* = the tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETR_t)$. *ATE* = aggregate earnings excluding the tax change component of earnings, measured as $PTE_t(1 - ETR_{t-1})$. *INTCC* = the initial tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETRQ1_t)$. *REVTCC* = the revised tax change component of earnings, measured as $PTE_t(ETRQ1_t - ETR_t)$. All regression variables (except *AR*), are scaled by contemporaneous average total assets.

[c] Panel C variables are defined as follows: *MVE* = market value of equity (Compustat annual item #25 * #199); *BMR* = book-to-market ratio (Compustat annual item #60 ÷ *MVE*); *PBHR* = 12-month buy-and-hold return, defined above; *BETA* = systematic risk estimated from regressions of monthly raw returns on a value-weighted portfolio over a 60-month period prior to the abnormal return accumulation period; $SPEC_t$ = special items (Compustat annual item #17); *DTAX* = deferred income tax expense (Compustat annual item #50); *VOL* = annual trading volume divided by common shares outstanding; *EP* = the earnings-to-price ratio, measured as income before extraordinary items divided by the fiscal year-end stock price; *PTXACC* = pretax accruals, defined as pretax income minus pretax operating cash flows; and *PTXCF* = pretax operating cash flow, measured as operating cash flow plus cash taxes paid minus cash flow due to extraordinary items (Compustat annual item #308 + #317 − #124). *SPEC*, *DTAX*, *PTXACC*, and *PTXCF* are scaled by contemporaneous average total assets.

[d] All means are multiplied by 10,000.

(t = 4.7462 and 2.4925, respectively). I obtain similar inferences using the nonparametric Wilcoxon signed-rank test ($ETRQ1_t$ is significantly higher than ETR_{t-1} and ETR_t [z = 19.102 and 14.530, respectively]). This seems to indicate that on average, management sets the estimate of the annual ETR too high in Q1, perhaps to create a reserve from which current earnings can be drawn, or because tax departments are conservative in their initial estimates (Comprix et al. 2006).[24]

Panel B of Table 1 reports the descriptive statistics for the variables used in the regression analyses. The mean of E_{t+1} is 0.0659, which indicates the average scaled earnings for the sample firms is approximately 6.6 percent. The average TCC_t accounts for approximately 2.1 percent of the average change in return on assets (ROA). $INTCC_t$ decreases ROA by approximately 5 percent on average, while $REVTCC_t$ increases ROA by approximately 3 percent on average. Therefore, for the average firm in the sample, approximately 60 percent of $INTCC_t$ reverses by year-end. These statistics are consistent with those from Panel A and quantifies the effect of tax changes on earnings. Panel C presents descriptive statistics for the risk and anomaly control variables.

---

[24] The unique feature of the integral view of quarterly reporting for taxes is that the estimation of the initial ETR creates opportunities for intra-year earnings management (annual ETR estimates are used to calculate quarterly income tax expense), which would affect current annual earnings, but not necessarily future annual earnings.

Panel D of Table 1 presents industry descriptive statistics, where the industry breakdown is adapted from Barth et al. (1999) and Rajgopal et al. (2003). The majority of the sample is clustered in four industries: financial institutions (20.17 percent), durable manufacturers (19.08 percent), retail (12.24 percent), and computers (7.63 percent). I also examine the means of TCC_t, $INTCC_t$, and $REVTCC_t$ at the industry level in Panel D of Table 1. The degree of reversal in $INTCC_t$ by the end of the year is startling for some industries. For example, in the food industry, 159 percent of $INTCC_t$ reverses by year-end, while 101 percent of $INTCC_t$ reverses by year-end for firms in the chemicals industry. Other industries exhibiting this strong reversal in $INTCC_t$ include textiles, printing and publishing (97 percent), retail (67 percent), pharmaceuticals (95 percent), and durable manufacturers (75 percent).

Table 2 reports Pearson correlations. Consistent with Guenther and Jones (2003), TCC_t is positively correlated with contemporaneous returns ($\rho_P = 0.0218$, $p = 0.0140$). $INTCC_t$ is also positively correlated with contemporaneous returns, while the correlation between $REVTCC_t$ and returns is not significant. $INTCC_t$ ($REVTCC_t$) is negatively (positively) correlated with ATE_t ($\rho_P = -0.1029$, $p = 0.0000$ and $\rho_P = 0.0270$, $p = 0.0024$). Finally, the correlation between $INTCC_t$ and $REVTCC_t$ is significantly negative ($\rho_P = -0.2778$, $p = 0.0000$), as is the correlation between pretax accruals and pretax operating cash flow ($\rho_P = -0.6455$, $p = 0.0000$).

Multivariable Results

Tables 3 and 4 present the results from estimating Equations (2) and (3) using OLS.[25] Because the data are pooled, I estimate the regressions in the spirit of Fama and MacBeth (1973). I estimate yearly cross-sectional regressions and use the average slopes and their time-series standard errors to draw inferences.[26] In Tables 3 and 4, I list the mean coefficient estimate first, followed by the Fama-MacBeth t-statistic and the number of yearly positive (negative) coefficients.

I also partition the sample and estimate Equations (2) and (3) separately for firms with decreases (increases) in annual ETRs. Because decreases in annual ETRs increase earnings (or conversely, increases in annual ETRs decrease earnings), *ceteris paribus*, managers will

---

[25] Results from the Breusch and Pagan (1980) Lagrange Multiplier Test ($\chi^2 = 0.59$, $p = 0.4431$) indicate that a fixed or random effects specification is not appropriate for Equation (2). This result holds for all specifications of Equations (2) and (3).

[26] *Assuming independence through time*, the Fama-MacBeth regressions control for cross-sectional correlation in the residuals. In the presence of positive cross-correlation in the residuals, the standard errors can be biased downward and the t-statistics biased upward in pooled regressions. For example, Fama and French (2000) find that Fama-MacBeth standard errors are two to five times the OLS standard errors from pooled regressions using their sample.

However, the assumption of temporal independence in regressions without abnormal returns as the dependent variable is tenuous, as many firms appear in the regressions multiple years. Any autocorrelation in the coefficients would further bias the standard errors, but reliably estimating the serial correlation from eight observations is problematic. Fama and French (2000) suggest that if the first-order autocorrelation is 0.5, t-statistics of 2.8 (instead of 1.96) should be used to infer reliability. Accordingly, I require t-statistics of 2.8 to indicate significance of the coefficient estimates in Equations (2), (3), (6), and (7).

TABLE 2
Sample Correlations

	ATE$_t$	*TCC$_t$*	*INTCC$_t$*	*REVTCC$_t$*	*MVE$_t$*	*BMR$_t$*	*PBHR$_t$*	*BETA$_t$*	*SPEC$_t$*	*DTAX$_t$*	*VOL$_t$*	*EP$_t$*	*PTXACC$_t$*	*PTXCF$_t$*
ATE$_t$														
TCC$_t$	**−0.0799** **(0.0000)**													
INTCC$_t$	**−0.1029** **(0.0000)**	**0.7665** **(0.0000)**												
REVTCC$_t$	**0.0270** **(0.0024)**	**0.4040** **(0.0000)**	**−0.2778** **(0.0000)**											
MVE$_t$	**0.0623** **(0.0000)**	0.0126 (0.1548)	**0.0219** **(0.0136)**	−0.0123 (0.1654)										
BMR$_t$	−0.0083 (0.3508)	−0.0098 (0.2689)	−0.0156 (0.0785)	0.0076 (0.3939)	−0.0103 (0.2447)									
PBHR$_t$	**−0.0175** **(0.0492)**	**0.0218** **(0.0140)**	**0.0180** **(0.0421)**	0.0069 (0.4360)	**−0.0247** **(0.0055)**	**0.0371** **(0.0000)**								
BETA$_t$	**0.1688** **(0.0000)**	**−0.0224** **(0.0255)**	−0.0120 (0.2304)	−0.0165 (0.1000)	0.0109 (0.2789)	**−0.0620** **(0.0000)**	**0.0555** **(0.0000)**							
SPEC$_t$	**0.0805** **(0.0000)**	**0.0187** **(0.0445)**	**−0.0310** **(0.0008)**	**0.0724** **(0.0000)**	**−0.0279** **(0.0026)**	−0.0039 (0.6749)	−0.0110 (0.2382)	**−0.0507** (0.0000)						
DTAX$_t$	**0.0505** **(0.0000)**	0.0166 (0.1007)	0.0189 (0.0612)	−0.0021 (0.8388)	**0.0220** **(0.0289)**	**0.0356** **(0.0004)**	**0.0385** **(0.0000)**	**0.0726** **(0.0000)**	**−0.2229** **(0.0000)**					
VOL$_t$	**0.3029** **(0.0000)**	**−0.0197** **(0.0274)**	**−0.0197** **(0.0269)**	−0.0013 (0.8821)	−0.0110 (0.2161)	**−0.0239** **(0.0074)**	**0.0238** **(0.0075)**	**0.4449** **(0.0000)**	**−0.0832** **(0.0000)**	**0.0911** **(0.0000)**				
EP$_t$	**−0.0331** **(0.0003)**	0.0163 (0.0771)	−0.0160 (0.0830)	**0.0475** **(0.0000)**	**−0.1125** **(0.0000)**	**0.0273** **(0.0031)**	**0.0720** **(0.0000)**	−0.1656 (0.0000)	**0.2484** (0.0000)	**−0.0710** (0.0000)	**−0.1384** (0.0000)			
PTXACC$_t$	**0.1201** **(0.0000)**	−0.0111 (0.2666)	**−0.0245** **(0.0143)**	0.0184 (0.0663)	**−0.0271** **(0.0000)**	−0.0031 (0.7579)	**−0.0623** **(0.0000)**	**0.1075** **(0.0000)**	**0.1794** **(0.0000)**	**−0.0561** **(0.0000)**	**0.1163** **(0.0000)**	**0.0824** **(0.0000)**		
PTXCF$_t$	**0.6643** **(0.0000)**	0.0058 (0.5647)	0.0031 (0.7563)	0.0042 (0.6717)	**0.0605** **(0.0000)**	**−0.0419** **(0.0000)**	**0.0328** **(0.0011)**	**0.0450** **(0.0000)**	**−0.0481** **(0.0000)**	**0.0807** **(0.0000)**	**0.1118** **(0.0000)**	**−0.0689** **(0.0000)**	**−0.6455** **(0.0000)**	

This table reports Pearson correlations for the full sample (n = 12,695). The correlation coefficient, followed by a two-sided p-value is reported in each cell. Bold values are significant at $p \leq 0.05$. Variables are defined as follows (all variables except *MVE*, *BMR*, *PBHR*, *VOL*, and *EP* are scaled by contemporaneous average total assets): *ATE* = aggregate earnings excluding the tax change, measured as *PTE$_t$*(1 − *ETR$_{t-1}$*). *TCC* = the tax change component of earnings, measured as *PTE$_t$*(*ETR$_{t-1}$* − *ETR$_t$*). *INTCC* = the initial tax change component of earnings, measured as *PTE$_t$*(*ETR$_{t-1}$* − *ETRQ1$_t$*). *REVTCC* = the revised tax change component of earnings, measured as *PTE$_t$*(*ETRQ1$_t$* − *ETR$_t$*). *MVE* = market value of equity. *BMR* = book to market ratio. *PBHR* = Prior 12-month buy-and-hold return. *BETA* = systematic risk estimated from regressions of monthly raw returns on a value-weighted portfolio over a 60-month period prior to the abnormal return accumulation period. *SPEC$_t$* = special items. *DTAX* = deferred income tax expense. *VOL* = annual trading volume divided by common shares outstanding. *EP* = the earnings-to-price ratio, measured as income before extraordinary items divided by the fiscal year-end stock price. *PTXACC* = pretax accruals, defined as pretax income minus pretax operating cash flows. *PTXCF* = pretax operating cash flow, measured as want to sustain ETR decreases and mitigate ETR increases. Therefore, the results also provide evidence on the persistence and forecasting implications of the tax change component of earnings under different conditions.

The Persistence and Forecasting Implications of the Tax Change Component of Earnings

Table 3 presents the results from estimating Equation (2), in which future earnings are regressed on current earnings and the tax change component. The estimate of the persistence of ATE_t in Panel A is 0.8106 (t = 34.53), while the persistence estimate of TCC_t is 0.7121 (t = 6.56). The coefficient on TCC_t is clearly different from zero, which is initial evidence that the tax change component of earnings is not transitory. When I estimate Equation (2) separately for firms with either annual ETR decreases (Panel B) or increases (Panel C), the results continue to indicate that the tax change component is not transitory. Specifically,

TABLE 3

Summary Regression Statistics of the Forecasting Implications of the Tax Change Component of Earnings[a]

$$E_{t+1} = \gamma_0 + \gamma_1 ATE_t + \gamma_2 TCC_t + \nu_{t+1} \tag{2}$$

Panel A: Full Sample (n = 8)

Parameters[b]	γ_0	γ_1	γ_2	$\bar{R}^2$
Estimate	0.0091	0.8106	0.7121	0.6041
t-statistic	9.49*	34.53*	6.56*[c]	37.80*
Pos/(Neg)	8/(0)	8/(0)	8/(0)	

Panel B: *ETR* Decreases (n = 8)

Parameters	γ_0	γ_1	γ_2	$\bar{R}^2$
Estimate	0.0073	0.8286	0.7373	0.6777
t-statistic	6.45*	35.02*	3.71*[c]	41.17*
Pos/(Neg)	8/(0)	8/(0)	7/(1)	

Panel C: *ETR* Increases (n = 8)

Parameters	γ_0	γ_1	γ_2	$\bar{R}^2$
Estimate	0.0108	0.7935	0.7198	0.5696
t-statistic	13.73*	39.36*	3.08*[c]	27.33*
Pos/(Neg)	8/(0)	8/(0)	6/(2)	

The symbols *, †, and ‡ denote conventional significance at the 0.01, 0.05, and 0.10 (two-tailed) levels, respectively. In order to account for potential serial correlation, I interpret t-statistics of 2.8 or higher on the regression coefficients as an indication of significance at the 0.05 level (Fama and French 2000).

[a] This table presents summary statistics of annual cross-sectional regressions estimated using the Fama and MacBeth (1973) approach. I use average slopes and their time-series standard errors to draw inferences (mean coefficients are reported first, followed by the t-statistic and the number of yearly positive (negative) coefficients). OLS is used rather than a fixed or random effects model. Results from Breusch and Pagan (1980) LM tests indicate that OLS is the appropriate estimation method.

[b] Variables are defined as follows (all variables are scaled by average total assets): *E* = income before extraordinary items; *ATE* = aggregate earnings excluding the tax change, measured as $PTE_t(1 - ETR_{t-1})$; and *TCC* = the tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETR_t)$.

[c] RQ_1: Is the tax change component of earnings useful in forecasting future earnings incremental to earnings excluding the tax change?

when annual ETRs decrease (increase), the coefficient on TCC_t is 0.7373 (0.7198) with a t-statistic of 3.71 (3.08). These results contradict the maintained hypothesis in prior literature that earnings generated by ETR changes are transitory.

If the earnings innovation to TCC_t were completely persistent, then γ_2 in Table 3 would be 1. However, the magnitude of the coefficient on TCC_t in each specification of Equation (2) indicates that TCC_t contains transitory elements. The results in the next section should help to identify the origins of the persistence of the tax change component of earnings.

The Persistence and Forecasting Implications of the Tax Change Component of Earnings in Interim Periods

Table 4 reports the estimation results for Equation (3), in which future earnings are regressed on current earnings, the initial tax change component, and the revised tax change component.[27] The coefficient on ATE_t in Panel A is 0.8137 ($t = 34.02$). The coefficient on $INTCC_t$ is 0.9398 ($t = 7.50$) and the coefficient on $REVTCC_t$ is 0.1116 ($t = 0.72$). The test of H_1 rejects the hypothesis that $\delta_2 = \delta_3$ ($t = 4.89$). Therefore, for the full sample, the initial tax change component of earnings is more persistent for future earnings than the revised tax change component of earnings.

Panel B of Table 4 reports the results of estimating Equation (3) for firms with annual ETR decreases. The coefficient on $INTCC_t$ is 0.9627 ($t = 3.65$) and the coefficient on $REVTCC_t$ is 0.2909 ($t = 2.20$). The test of H_1 rejects the hypothesis that $\delta_2 = \delta_3$ ($t = 2.53$). Results from estimating Equation (3) for firms with annual ETR increases (Panel C) show that $INTCC_t$ is also more persistent for future earnings than $REVTCC_t$ ($\delta_2 = 0.9144$, $t = 4.29$; $\delta_3 = -0.0505$, $t = -0.15$). The test of H_1 rejects the hypothesis that $\delta_2 = \delta_3$ ($t = 6.09$).

Prior literature has generally found that the coefficient on pretax earnings is greater than that on ETR-related earnings. Although this is not the focus of my study, unreported tests indicate that the difference in the persistence of the tax (*TCC* and *INTCC*) and earnings coefficients (*ATE*) is not statistically significant. The failure to reject the null hypothesis of no difference could be seen as further evidence against the hypothesis that ETR-related earnings are transitory.

When I decompose $REVTCC_t$ to examine the forecasting implications of the tax change component for each individual quarter (Q2, Q3, and Q4), the results indicate that the tax change component becomes more transitory (less persistent) as the year progresses (i.e., Q1 > Q2 > Q3 > Q4). The coefficient estimates for TCC_t in each quarter are: Q1 = 0.9546; Q2 = 0.4836; Q3 = 0.0937; Q4 = −0.1838. Both Q1 and Q2 are significant at $p < 0.05$.

I also find that δ_2 is greater than zero and δ_3 is not statistically different from zero in all reported specifications of Equation (3). These results suggest that, on average, management's initial ETR estimate reflects the persistent information contained in the annual tax change.[28] By decomposing TCC_t into an initial and revised portion, I identify a persistent

---

[27] Results of a series of F-tests indicate that the decomposition of earnings into these three components is more informative than a univariate regression of future earnings on current earnings ($F = 81.61$, $p = 0.0000$) or a regression of future earnings on current earnings and the aggregate tax change component ($F = 56.08$, $p = 0.0000$).

[28] The phrase "on average" should be emphasized. If I partition the sample into firms that have increases in *INTCC* and decreases in *REVTCC*, then the persistence estimates of *TCC* (in Equation (2)) and *INTCC* (in Equation (3)) are not significantly different from zero. However, *INTCC* is still significantly larger that *REVTCC* in this situation.

TABLE 4
Summary Regression Statistics of the Forecasting Implications of the Initial and Revised Tax Change Component of Earnings[a]

$$E_{t+1} = \delta_0 + \delta_1 ATE_t + \delta_2 INTCC_t + \delta_3 REVTCC_t + \nu_{t+1} \tag{3}$$

Panel A: Full Sample (n = 8)

Parameters[b]	δ_0	δ_1	δ_2	δ_3	$\bar{R}^2$	H_1: $\delta_2 > \delta_3$[c]
Estimate	0.0092	0.8137	0.9398	0.1116	0.6562	0.8281
t-statistic	9.33*	34.02*	7.50*	0.72	37.12*	4.89*
Pos/(Neg)	8/(0)	8/(0)	8/(0)	6/(2)		7(1)

Panel B: *ETR* Decreases (n = 8)

Parameters	δ_0	δ_1	δ_2	δ_3	$\bar{R}^2$	H_1: $\delta_2 > \delta_3$
Estimate	0.0076	0.8290	0.9627	0.2909	0.6997	0.6718
t-statistic	6.42*	34.15*	3.65*	2.20	41.50*	2.53†
Pos/(Neg)	8/(0)	8/(0)	7/(1)	5/(3)		7(1)

Panel C: *ETR* Increases (n = 8)

Parameters	δ_0	δ_1	δ_2	δ_3	$\bar{R}^2$	H_1: $\delta_2 > \delta_3$
Estimate	0.0106	0.7958	0.9144	−0.0505	0.6225	0.9649
t-statistic	13.20*	38.60*	4.29*	−0.15	26.74*	6.09*
Pos/(Neg)	8/(0)	8/(0)	8/(0)	5/(3)		8/(0)

The symbols *, †, and ‡ denote conventional significance at the 0.01, 0.05, and 0.10 (two-tailed) levels, respectively. In order to account for potential serial correlation, I interpret t-statistics of 2.8 or higher on the regression coefficients as an indication of significance at the 0.05 level (Fama and French 2000). The symbols for tests of H_1 represent significance levels for one-tailed tests.

[a] This table presents summary statistics of annual cross-sectional regressions estimated using the Fama and MacBeth (1973) approach. I use average slopes and their time-series standard errors to draw inferences (mean coefficients are reported first followed by the *t*-statistic and the number of yearly positive (negative) coefficients). OLS is used rather than a fixed or random effects model. Results from Breusch and Pagan (1980) LM tests indicate that OLS is the appropriate estimation method.

[b] Variables are defined as follows (all variables are scaled by average total assets): E = income before extraordinary items; ATE = aggregate earnings excluding the tax change, measured as $PTE_t(1 - ETR_{t-1})$; $INTCC$ = the initial tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETRQ1_t)$; and $REVTCC$ = the revised tax change component of earnings, measured as $PTE_t(ETRQ1_t - ETR_t)$.

[c] H_1: The initial (Q1) tax change component of earnings is more persistent for future earnings than subsequent revisions (Q2, Q3, Q4) in the tax change component of earnings.

element of the tax change component of earnings, which further dispels the notion that earnings generated by ETR changes are transitory.

Valuation Implications of the Tax Change Component of Earnings

Panel A of Table 5 reports results from estimating the system of Equations (2) and (4) on the full sample. For ATE_t, the valuation coefficient ($\gamma_1^* = 0.9131$) is larger than the forecasting coefficient ($\gamma_1 = 0.8149$), which suggests that the market overestimates the persistence of earnings excluding the tax change relative to its ability to forecast one-year-ahead earnings. The likelihood ratio statistic reported in Panel A is statistically significant ($\chi^2(1) = 6.3403$, $p = 0.009$). The market appears to underestimate the persistence of TCC_t;

TABLE 5
Nonlinear Generalized Least Squares Estimation of the Pricing of Earnings Components[a]

Panel A: Market Pricing of the Tax Change Component with Respect to its Implications for One-Year-Ahead Earnings

$$E_{t+1} = \gamma_0 + \gamma_1 ATE_t + \gamma_2 TCC_t + \nu_{t+1} \quad (2)$$

$$AR_{t+1} = \phi_0 + \phi_1(E_{t+1} - \gamma_0 - \gamma_1^* ATE_t - \gamma_2^* TCC_t) + \varepsilon_{t+1} \quad (4)$$

Parameters[b]	γ_1	γ_2	γ_1^*	γ_2^*	
Estimate	0.8149	0.6914	0.9131	−0.2276	
Asymptotic SE	0.0055	0.0669	0.0391	0.4707	
Test of Market Efficiency:[c]			$\gamma_1 = \gamma_1^*$	$\gamma_2 = \gamma_2^*$	$\gamma_k = \gamma_k^*$
Likelihood Ratio Statistic:			6.3403*	3.7913*	10.9818*

Panel B: Market Pricing of the Initial and Revised Tax Change Component with Respect to their Implications for One-Year-Ahead Earnings

$$E_{t+1} = \delta_0 + \delta_1 ATE_t + \delta_2 INTCC_t + \delta_3 REVTCC_t + \nu_{t+1} \quad (3)$$

$$AR_{t+1} = \omega_0 + \omega_1(E_{t+1} - \delta_0 - \delta_1^* ATE_t - \delta_2^* INTCC_t - \delta_3^* REVTCC_t) + \varepsilon_{t+1} \quad (5)$$

Parameters	δ_1	δ_2	δ_3	δ_1^*	δ_2^*	δ_3^*	
Estimate	0.8176	0.9147	0.0977	0.9179	0.2571	−1.4850	
Asymptotic SE	0.0055	0.0731	0.1036	0.0388	0.5080	0.7243	
Test of Market Efficiency:[d]				$\delta_1 = \delta_1^*$	$\delta_2 = \delta_2^*$	$\delta_3 = \delta_3^*$	$\delta_k = \delta_k^*$
Likelihood Ratio Statistic:				6.7291*	1.6526	4.8654*	12.4737*

The symbols *, †, and ‡ denote significance at the 0.01, 0.05, and 0.10 (two-tailed) levels, respectively.

[a] This table reports the estimation results of the Mishkin (1983) framework. The two systems of equations are jointly estimated using an iterative nonlinear generalized least squares estimation procedure based on 12,695 observations during 1994–2001.

[b] Variables are defined as follows (all variables are scaled by average total assets): E = income before extraordinary items; ATE = aggregate earnings excluding the tax change, measured as $PTE_t(1 - ETR_{t-1})$; TCC = the tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETR_t)$; $INTCC$ = the initial tax change component of earnings, measured as $PTE_t(ETR_{t-1} - ETRQ1_t)$; and $REVTCC$ = the revised tax change component of earnings, measured as $PTE_t(ETRQ1_t - ETR_t)$.

[c] H_{2a}: The earnings expectations embedded in stock prices fail to fully reflect the persistence of the tax change component of earnings.

[d] H_{2b}: The earnings expectations embedded in stock prices fail to fully reflect the persistence of the initial (Q1) and revised (Q2, Q3, Q4) tax change component of earnings.

the valuation coefficient ($\gamma_2^* = -0.2276$) is significantly smaller than the forecasting coefficient ($\gamma_2 = 0.6914$; $\chi^2(1) = 3.7913$, $p = 0.048$).

Panel B of Table 5 reports results from the estimation of the system of Equations (3) and (5) on the full sample. The results for ATE_t are similar to those reported in Panel A: the valuation coefficient ($\delta_1^* = 0.9179$) is significantly larger than the forecasting coefficient ($\delta_1 = 0.8176$), which suggests that the market overestimates the persistence of earnings relative to its ability to forecast one-year-ahead earnings. The market appears to underestimate both the persistence of $INTCC_t$ and the persistence of $REVTCC_t$. Specifically, the valuation coefficient ($\delta_2^* = 0.2571$) is smaller than the forecasting coefficient ($\delta_2 = 0.9147$)

for *INTCC*$_t$, and the valuation coefficient ($\delta_3^* = -1.4850$) is less than the forecasting coefficient ($\delta_3 = 0.0977$) for *REVTCC*$_t$. The likelihood ratio statistics reported in Panel B indicate that the mispricing of *INTCC*$_t$ is not statistically significant ($\chi^2(1) = 1.652$, p = 0.199), while the mispricing of *REVTCC*$_t$ is statistically significant ($\chi^2(1) = 4.865$ p = 0.027).[29] These results suggest that investors do not fully understand the implications of the transitory nature of the revised tax change component.

The lower valuation coefficients (relative to the forecasting coefficients) on *TCC* and *REVTCC* are consistent with a market that punishes firms for engaging in behavior it deems representative of earnings management. DeFond and Park (2001), Bartov et al. (2002), and Gleason and Mills (2004) find that the market discounts the reward to beating earnings expectations when earnings appear to be managed.

Robustness Checks and Sensitivity Analyses

Tax (or Benefit) Applicable to Significant Unusual or Infrequently Occurring Items, Discontinued Operations, or Extraordinary Items

A provision of APB No. 28 excludes taxes related to "significant unusual or extraordinary items that will be separately reported or reported net of their related tax effect" from the estimated annual ETR calculation (APB No. 28, ¶19). Instead, these items should be recognized in the interim period in which they occur (APB No. 28, ¶21). To the extent that these significant unusual or extraordinary items are not randomly distributed throughout the year, they could drive the results for H_1. A similar concern exists with respect to in-process research and development (IPRD), a large ETR reconciliation item that is transitory. Because IPRD is not persistent, the transitory nature of *REVTCC* could be due to IPRD (instead of earnings management) to the extent that IPRD is either not predicted well in Q1 or IPRD is booked disproportionately in Q2–Q4. To address this concern, I identify firms that reported a merger or acquisition, divestiture, special item, or extraordinary item (Compustat item numbers FN1, 66, 17, and 192, respectively), exclude them from the sample (leaving a sample of 10,375 firm-year observations), and rerun the tests for H_1. Untabulated results indicate that the null continues to be rejected for each hypothesis.

Abnormal Returns to a Hedge Strategy

The findings from the Mishkin (1983) test suggest that stock prices do not accurately reflect the persistence of the tax change component of earnings and this mispricing is driven by the revised tax change component of earnings. In order to control for additional risk and anomaly variables, I supplement the Mishkin (1983) tests by estimating return regressions using the Fama and MacBeth (1973) technique.

The Fama and MacBeth (1973) approach calls for the formation of zero-investment hedge portfolios. To create long and short positions of the hedge portfolios, I first rank firms on the magnitude of *ATE*$_t$ and equally assign firms to one of ten portfolios each year. Next, within each of the ten *ATE* portfolios, I assign firms equally into deciles based on *INTCC*$_t$ and *REVTCC*$_t$. I divide each *INTCC*$_t$ and *REVTCC*$_t$ decile number by 9 so each observation related to *INTCC*$_t$ or *REVTCC*$_t$ takes on a value ranging from 0 (for the lowest decile) to 1 (for the highest decile). I form the portfolios in this manner in order to rank firms on *INTCC*$_t$ or *REVTCC*$_t$ while controlling for the level of *ATE*$_t$. Firms in the lowest decile are firms with highest values of *INTCC*$_t$ and *REVTCC*$_t$ for a given level of *ATE*$_t$, while firms in the highest decile have the lowest values of $INTCC_t$ and $REVTCC_t$ for a given level of ATE_t. The hedge portfolio is formed by taking a long (short) position in firms in the highest (lowest) decile of $INTCC_t$ or $REVTCC_t$.

---

[29] Because the data used in the Mishkin estimation are pooled, I also estimate the system of Equations using the Fama and MacBeth (1973) approach, the benefits of which are discussed in footnote 26. The coefficient estimates from each equation are qualitatively similar, and a two-tailed difference of means t-test of the hypothesis that the mean valuation and forecasting coefficients of *REVTCC* are equal is rejected (t = 2.38, p = 0.046).

I implement the Fama and MacBeth (1973) approach by estimating the following cross-sectional regression for each of the eight years in the sample:

$$AR_{t+1} = \varphi_0 + \varphi_1 INTCC_t^{dec} + \varphi_2 REVTCC_t^{dec} + \xi_{t+1} \tag{6}$$

where $INTCC_t^{dec}$ represents the scaled portfolio decile rank of the initial tax change component of earnings and $REVTCC_t^{dec}$ represents the scaled portfolio decile rank of the revised tax change component of earnings. In order for the regression coefficients to be interpreted as zero-investment portfolio returns, the returns used in the regressions must be for the same time period and the accounting information used to create the hedge portfolios must coincide chronologically and be publicly available at the time of portfolio formation. Thus, I estimate Equation (6) using firms with December year-ends only, which reduces my sample size to 8,061 firm-year observations. To further insure that the information is available to the market during the return calculation period, I also require firms to have Compustat earnings announcement dates for all four quarters in the subsequent year and these dates must fall within the 12-month buy-and-hold return interval (Thomas 2000).

The results from the Fama and MacBeth (1973) regressions in Panel A of Table 6 generally confirm the results of the Mishkin (1983) test. There is an insignificant, positive association between $INTCC_t$ and future returns in Equation (6). There is also a positive association between $REVTCC_t$ and future returns in Equation (6), however the association is marginally significant (φ_2 = 0.0487, t = 2.09). The positive sign on $REVTCC_t$ is consistent with the difference in historical and security market weightings documented using the Mishkin (1983) framework. Because the market underestimates the persistence of $REVTCC_t$, there should be positive abnormal returns for portfolios ranked on $REVTCC_t$. Thus, the abnormal return to a trading strategy based on $REVTCC_t$ is 4.87 percent.

Controlling for Additional Risk Factors and Market Anomalies

Although my measure of abnormal returns controls for three of the four commonly identified risk factors (Fama and French 1993; Carhart 1997), I also estimate the following regression to ensure that any relation between the revised tax change component of earnings and future abnormal returns is incremental to other risk factors and market anomalies identified in prior research:

$$\begin{aligned} AR_{t+1} = {} & \varphi_0 + \varphi_1 INTCC_t^{dec} + \varphi_2 REVTCC_t^{dec} + \varphi_3 BETA_t^{dec} + \varphi_4 SPEC_t^{dec} \\ & + \varphi_5 DTAX_t^{dec} + \varphi_6 VOL_t^{dec} + \varphi_7 EP_t^{dec} + \varphi_8 PTXCF_t^{dec} \\ & + \varphi_9 PTXACC_t^{dec} + \xi_{t+1} \end{aligned} \tag{7}$$

where $BETA_t$ = the slope coefficient from a regression of raw monthly returns on the value-weighted portfolio return over a 60-month period prior to the abnormal return accumulation period (Fama and French 1992), $SPEC_t$ = special items (annual Compustat item #17), computed as special items divided by average total assets (Elliott and Hanna 1996; Easton et al. 2000; Burgstahler et al. 2002), $DTAX_t$ = deferred income tax expense (annual Compustat item #50) divided by average total assets (Chaney and Jeter 1994; Philips et al. 2003; Joos et al. 2003; Hanlon 2005), VOL_t = the firm-specific annual trading volume from CRSP divided by common shares outstanding (Barth and Hutton 2004), EP_t = the earnings to

TABLE 6

Summary Regression Statistics of the Relation between Future Abnormal Stock Returns and Scaled Initial and Revised Tax Change Component Decile Rankings (Fama and MacBeth [1973] Approach)[a]

$$AR_{t+1} = \varphi_0 + \varphi_1 INTCC_t^{dec} + \varphi_2 REVTCC_t^{dec} + \xi_{t+1} \tag{6}$$

$$AR_{t+1} = \varphi_0 + \varphi_1 INTCC_t^{dec} + \varphi_2 REVTCC_t^{dec} + \varphi_3 BETA_t^{dec} + \varphi_4 SPEC_t^{dec} + \varphi_5 DTAX_t^{dec} + \varphi_6 VOL_t^{dec} + \varphi_7 EP_t^{dec} + \varphi_8 PTXCF_t^{dec} + \varphi_9 PTXACC_t^{dec} + \xi_{t+1} \tag{7}$$

Panel A: Univariate Regression—Total Period Return (n = 8,061)

Parameters[b]	φ_0	φ_1	φ_2	$\bar{R}^2$
Estimate	0.0334	0.0024	0.0487	0.0032
t-statistic	0.40	0.15	2.09‡	2.40‡
Pos/(Neg)	3/(5)	4/(4)	6/(2)	

Panel B: Regression after Controlling for Additional Risk Factors and Market Anomalies (n = 5,828)

Parameters	φ_0	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6	φ_7	φ_8	φ_9	$\bar{R}^2$
Estimate	−0.0114	−0.0028	0.0594	0.1351	−0.0320	0.0425	0.0071	0.0757	−0.0400	−0.1234	0.0693
t-statistic	−0.09	−0.22	2.55†	0.96	−1.56	1.59	0.14	1.18	0.92	−2.83*	3.56*
Pos/(Neg)	4/(4)	2/(6)	7/(1)	6/(2)	2/(6)	6/(2)	4/(4)	5/(3)	2/(6)	2/(6)	

The symbols *, †, and ‡ denote conventional significance at the 0.01, 0.05, and 0.10 (two-tailed) levels, respectively. In order to account for potential serial correlation, I interpret t-statistics of 2.8 or higher on the regression coefficients as an indication of significance at the 0.05 level (Fama and French 2000).

[a] This table presents summary statistics of eight annual cross-sectional regressions estimated using the Fama and MacBeth (1973) approach. I use average slopes and their time-series standard errors to draw inferences (mean coefficients are reported first, followed by the t-statistic and the number of yearly positive (negative) coefficients). The decile rankings of *INTCC* and *REVTCC* are formed as follows: Ten portfolios are formed by first ranking firms on the magnitude of after-tax earnings, excluding the tax change (*ATE*). Next, within each of the ten *ATE* portfolios, firms are sorted evenly into deciles based on the initial tax change component of earnings and the revised tax change component of earnings. The long position consists of firms within the highest decile of each portfolio (i.e., *INTCC* or *REVTCC*), and the short position consists of firms in the lowest decile of each portfolio (i.e., *INTCC* or *REVTCC*). I estimate Equation (6) and Equation (7) using firms with December year-ends only, which results in a sample size of 8,061 and 5,828 firm-year observations, respectively.

[b] Variables are defined as follows: AR = Abnormal returns are defined as the 12-month buy-and-hold security return beginning April 1 and ending March 31 minus the value-weighted return of a benchmark size/book-to-market/prior year return portfolio over the same period (See Table 1 for a description of the benchmark portfolio construction); $INTCC^{dec}$ = the initial tax change component of earnings, transformed into a scaled-decile variable with values ranging from 0 to 1; $REVTCC^{dec}$ = the revised tax change component of earnings, transformed into a scaled-decile variable with values ranging from 0 to 1; $BETA^{dec}$ = systematic risk estimated from regressions of monthly raw returns on a value-weighted portfolio over a 60-month period prior to the abnormal return accumulation period, transformed into a scaled-decile variable with values ranging from 0 to 1; $SPEC^{dec}$ = special items, computed as special items divided by average total assets, transformed into a scaled-decile variable with values ranging from 0 to 1; $DTAX^{dec}$ = deferred income tax expense divided by average total assets, transformed into a scaled-decile variable with values ranging from 0 to 1; VOL^{dec} = annual trading volume divided by common shares outstanding, transformed into a scaled-decile variable with values ranging from 0 to 1; EP^{dec} = the earnings-to-price ratio, measured as income before extraordinary items divided by the fiscal year-end stock price, transformed into a scaled-decile variable with values ranging from 0 to 1; $PTXCF^{dec}$ = total pretax operating cash flow, measured as operating cash flows minus cash flow due to extraordinary items plus taxes paid in cash divided by average total assets, transformed into a scaled-decile variable with values ranging from 0 to 1; and $PTXACC^{dec}$ = total pretax operating accruals, measured as pretax income minus pretax operating cash flows divided by average total assets, transformed into a scaled-decile variable with values ranging from 0 to 1.

price ratio (Basu 1977), *PTXCF*$_t$ = total pretax operating cash flow, measured as operating cash flow (annual Compustat item #308) minus cash flow due to extraordinary items (annual Compustat data item #124) plus taxes paid in cash (annual Compustat data item #317) divided by average total assets (Hanlon 2005), *PTXACC*$_t$ = total pretax operating accruals, measured as pretax income minus pretax operating cash flows divided by average total assets (Sloan 1996; Collins and Hribar 2002). All risk and market anomaly controls are transformed into scaled decile variable with values ranging from 0 to 1. The inclusion of the additional control variables reduces the sample size used to estimate Equation (7) to 5,828 observations. The results from the Fama and MacBeth (1973) regressions in Panel B of Table 6 indicate that the marginally significant, incremental abnormal returns related to *REVTCC*$_t$ persist after controlling for additional risk factors and market anomalies ($\varphi_1 = 0.0594$, $t = 2.55$).

VI. CONCLUSIONS

I examine whether earnings generated by changes in ETRs (the tax change component) persist and aid in forecasting future earnings incremental to aggregate earnings excluding the tax change. Additionally, this study extends prior literature by testing whether the market fully incorporates the pricing effects of the persistence of the tax change component of earnings.

I find that there is a positive, significant association between the tax change component of earnings and future earnings. When I decompose the tax change component into an initial and revised portion, I find that the initial tax change component is more persistent and, therefore, more useful in forecasting future earnings than the revised tax change component. These results are consistent with my hypotheses that the initial and revised tax change components have differential persistence and forecasting implications, and dispute the broad notion that ETR-related earnings are transitory. Results from market tests indicate that investors underweight the forecasting implications of the tax change component and the mispricing appears to be driven by the transitory nature of the revised tax change component. Supplemental tests confirm that the results obtained from the market tests are likely a result of mispricing rather than alternative explanations such as omitted risk factors.

The results also suggest that investors can gain from breaking the tax change component of earnings into an initial and revised portion. To the extent that investors focus on the initial tax change component, they can better understand the persistence of earnings generated by ETR changes and, consequently, make better forecasts of future earnings. Such an understanding should also decrease the mispricing of ETR-related earnings. Investors may also benefit by focusing on the relative lack of persistence for the revised tax change component. When viewed in the context of concurrent research that finds Q2, Q3, and Q4 ETR revisions are used to meet or beat period-specific benchmarks (e.g., analyst forecasts), the lower persistence of the revised tax change component is consistent with managers using the revised tax change component as an earnings management tool. Increased attention to this issue may mitigate such behavior by managers.

REFERENCES

Abarbanell, J., and B. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35 (1): 1–24.

———, and ———. 1998. Abnormal returns to a fundamental analysis strategy. *The Accounting Review* 73 (1): 19–45.

Accounting Principles Board (APB). 1973. *Interim Financial Reporting*. Opinion No. 28. New York, NY: APB.

Amir, E., B. Lev, and T. Sougiannis. 2000. The incremental value of analysts' earnings forecasts in explaining stock returns. Working paper, Tel Aviv University.

Barth, M., W. Beaver, J. Hand, and W. Landsman. 1999. Accruals, cash flows, and equity values. *Review of Accounting Studies* 4 (3–4): 205–229.

———, ———, ———, and ———. 2004. Accruals, accounting-based valuation models, and the prediction of equity values. Working paper, Stanford University.

———, and A. Hutton. 2004. Analyst earnings forecast revisions and the pricing of accruals. *Review of Accounting Studies* 9 (1): 59–96.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (2): 173–204.

Basu, S. 1977. Investment performance of common stocks in relation to their price-earnings ratio: A test of the efficient markets hypothesis. *Journal of Finance* 32 (3): 663–682.

Bathke, A., and K. Lorek. 1984. The relationship between time-series models and the security market's expectation of quarterly earnings. *The Accounting Review* 59 (2): 163–176.

Bauman, C., M. Bauman, and R. Halsey. 2001. Do firms use the deferred tax asset valuation allowance to manage earnings? *The Journal of the American Taxation Association* 23 (Supplement): 27–48.

Bauman, M., and K. Shaw. 2005. Interim income tax data and earnings prediction. *The Journal of the American Taxation Association* 27 (2): 57–82.

Bodnar, G., and J. Weintrop. 1997. The valuation of the foreign income of multinational firms: A growth opportunities perspective. *Journal of Accounting and Economics* 24 (1): 69–97.

Boynton, C., and L. Mills. 2004. The evolving Schedule M-3: A new era of corporate show and tell? *National Tax Journal* 57 (3): 757–772.

Bradshaw, M., S. Richardson, and R. Sloan. 2001. Do analysts and auditors use information in accruals? *Journal of Accounting Research* 39 (1): 45–74.

Breusch, T., and A. Pagan. 1980. The Lagrange multiplier test and its application to model specification in econometrics. *Review of Economic Studies* 47 (1): 239–254.

Burgstahler, D., J. Jiambalvo, and T. Shevlin. 2002. Do stock prices fully reflect the implications of special items for future earnings? *Journal of Accounting Research* 40 (3): 585–612.

———, B. Elliott, and M. Hanlon. 2003. How firms avoid losses: Evidence of the use of the net deferred tax asset account. Working paper, University of Washington.

Carhart, M. 1997. On the persistence of mutual fund performance. *Journal of Finance* 52 (1): 57–82.

Chaney, P., and D. Jeter. 1994. The effect of deferred taxes on securities prices. *The Journal of Accounting, Auditing, and Finance* 9 (1): 91–116.

Chen, K., and M. Schoderbek. 2000. The 1993 tax rate increase and deferred tax adjustments: A test of functional fixation. *Journal of Accounting Research* 38 (1): 23–44.

Collins, D., and S. P. Kothari. 1989. An analysis of the intertemporal and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics* 11 (2–3): 143–181.

———, E. Maydew, and I. Weiss. 1997. Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics* 24 (1): 39–67.

———, and P. Hribar. 2002. Errors in estimating accruals: Implications for accounting research. *Journal of Accounting Research* 40 (1): 105–134.

Collins, W., W. Hopwood, and J. McKeown. 1984. The predictability of interim earnings over alternative quarters. *Journal of Accounting Research* 22 (2): 467–479.

Comprix, J., L. Mills, and A. Schmidt. 2006. Earnings management through changes in interim expenses. Working paper, Columbia University.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52 (3): 1035–1058.

DeFond, M., and C. Park. 2001. The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review* 76 (3): 375–404.

Dhaliwal, D., C. Gleason, and L. Mills. 2004. Last chance earnings management: Using the tax expense to achieve earnings targets. *Contemporary Accounting Research* 21 (2): 431–458.

Easton, P., P. Schroff, and G. Taylor. 2000. Permanent and transitory, accounting recording lag, and the earnings coefficient. *Review of Accounting Studies* 5 (4): 281–300.

Elgers, P., M. Lo, and R. Pfeiffer. 2001. Delayed security price adjustments to financial analysts' forecasts of annual earnings. *The Accounting Review* 76 (4): 613–632.

———, ———, and ———. 2003. Analysts vs. investors weightings of accruals in forecasting annual earnings. *Journal of Accounting and Public Policy* 22: 255–280.

Elliott, J., and J. Hanna. 1996. Repeated accounting write-offs and the information content of earnings. *Journal of Accounting Research* 34 (Supplement): 135–155.

Fairfield, P., S. Whisenant, and T. Yohn. 2003. The differential persistence of accruals and cash flows for future operating income versus future profitability. *Review of Accounting Studies* 8 (2): 221–243.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (May/June): 607–636.

———, and K. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (2): 427–465.

——— and ———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

——— and ———. 1998. Taxes, financing decisions, and equity values. *Journal of Finance* 53 (2): 819–843.

——— and ———. 2000. Forecasting profitability and earnings. *Journal of Business* 73 (2): 161–175.

Financial Accounting Standards Board (FASB). 1974. *Reporting Accounting Changes in Interim Financial Statements: An Amendment of APB Opinion No. 28.* Statement of Financial Accounting Standards No. 3. Stamford, CT: FASB.

———. 1977. *Accounting for Income Taxes in Interim Periods: An Interpretation of APB Opinion No. 28.* FASB Interpretation No. 18. Stamford, CT: FASB.

———. 1987. *Accounting for Income Taxes.* Statement of Financial Accounting Standards No. 96. Norwalk, CT: FASB.

———. 1992. *Accounting for Income Taxes.* Statement of Financial Accounting Standards No. 109. Norwalk, CT: FASB.

Frank, M., and S. Rego. 2006. Do managers use the valuation allowance account to manage earnings around certain earnings targets? *The Journal of the American Taxation Association* (forthcoming).

Gleason, C., and L. Mills. 2002. Materiality and contingent tax liability reporting. *The Accounting Review* 77 (2): 317–342.

———, and ———. 2004. Evidence of differing market responses to meeting or beating targets through tax expense management. Working paper, The University of Iowa.

Graham, J., and A. Tucker. 2005. Tax shelters and corporate debt policy. Working paper, Duke University.

Guenther, D., and D. Jones. 2003. Valuation implications of changes in firms' effective tax rates. Working paper, University of Colorado at Boulder.

Gupta, S., and K. Newberry. 1997. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting and Public Policy* 16 (1): 1–34.

———, and L. Mills. 2002. Corporate multistate tax planning: Benefits of multiple jurisdictions. *Journal of Accounting and Economics* 33 (1): 117–139.

Hanlon, M., and T. Shevlin. 2002. Accounting for tax benefits of employee stock options and implications for research. *Accounting Horizons* 16 (1): 1–16.

———. 2003. What can we infer about a firm's taxable income from its financial statements? *National Tax Journal* 56 (4): 831–863.

———. 2005. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review* 80 (1): 137–166.

———, S. Kelley, and T. Shevlin. 2005. Evidence on the possible information loss of conforming book income and taxable income. *Journal of Law and Economics* 48 (2): 407–442.

Hewlett Packard Company. 2000. Form 10-Q. July 31. Available at http://www.shareholder.com/Common/Edgar/47217/912057-00-41091/00-00.pdf.

Hribar, P. 2001. The market pricing of components of accruals. Working paper, Cornell University.

Joos, P., J. Pratt, and S. Young. 2003. Using deferred taxes to detect earnings management: Further evidence. Working paper, Massachusetts Institute of Technology.

Kothari, S. P., and J. Zimmerman. 1995. Price and return models. *Journal of Accounting and Economics* 20 (2): 155–192.

Kross, W., and D. Schroeder. 1990. An investigation of seasonality in stock price responses to quarterly earnings announcements. *Journal of Business, Finance, and Accounting* (Winter): 649–675.

Krull, L. 2004. Taxes and the reinvestment of foreign subsidiary earnings. *The Accounting Review* 79 (3): 745–768.

LaPorta, R. 1996. Expectations and the cross-section of stock returns. *Journal of Finance* 51 (5): 1715–1742.

Lev, B., and S. R. Thiagarajan. 1993. Fundamental information analysis. *Journal of Accounting Research* 31 (2): 190–215.

———, and D. Nissim. 2004. Taxable income, future, earnings, and equity values. *The Accounting Review* 79 (4): 1039–1074.

Lipe, R. 1986. The information contained in the components of earnings. *Journal of Accounting Research* 24 (1): 37–64.

———, and V. Bernard. 2000. Differences between interim and fourth-quarter earnings: An analysis of volatility and association with stock returns. Working paper, University of Oklahoma.

Livnat, J. 2003. Differential persistence of extremely negative and positive earnings surprises: Implications for the post-earnings announcement drift. Working paper, New York University.

McGill, G., and E. Outslay. 2004. Lost in translation: Detecting tax shelter activity in financial statements. *National Tax Journal* 57 (3): 739–756.

Mendenhall, R. 1991. Evidence on the possible underweighting of earnings-related information. *Journal of Accounting Research* 29 (1): 170–179.

Miller, G., and D. Skinner. 1998. Determinants of the valuation allowance for deferred tax assets under SFAS No. 109. *The Accounting Review* 73 (2): 213–233.

Mills, L., M. Erickson, and E. Maydew. 1998. Investments in tax planning. *The Journal of American Taxation Association* 20 (1): 1–20.

Mishkin, F. 1983. *A Rational Expectations Approach to Macroeconometrics: Testing Policy Effectiveness and Efficient Markets Models.* Chicago, IL: University of Chicago Press for the National Bureau of Economic Research.

Myers, J., L. Myers, and D. Skinner. 2005. Earnings management and earnings momentum. Working paper, Texas A&M University.

Nelson, K., J. Elliot, and R. Tarpley. 2003. How are earnings managed? Examples from auditors. *Accounting Horizons* 17 (Supplement): 17–35.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* 6 (1): 109–154.

O'Glove, T. 1987. *Quality of Earnings: The Investor's Guide to How Much Money a Company is Really Making.* New York, NY and London, U.K.: The Free Press.

Penman, S. 1992. Return to fundamentals. *Journal of Accounting, Auditing and Finance* 7 (3): 465–482.

Phillips, J., M. Pincus, and S. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521.

Rajgopal, S., T. Shevlin, and M. Venkatachalam. 2003. Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. *Review of Accounting Studies* 8 (4): 461–492.

Rangan, S., and R. Sloan. 1998. Implications of the integral approach to quarterly reporting for the post-earnings-announcement drift. *The Accounting Review* 73 (3): 353–371.

Reason, T. 2002. Align the books? *CFO Magazine*. November 1, 2002. CFO Publishing Corporation.

Salamon, G., and T. Stober. 1994. Cross-quarter differences in stock price responses to earnings announcements: Fourth quarter and seasonality influences. *Contemporary Accounting Research* 11 (1–2): 297–330.

Scholes, M., M. Wolfson, M. Erickson, E. Maydew, and T. Shevlin. 2004. *Taxes and Business Strategy: A Planning Approach*. 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Schrand, C., and M. H. F. Wong. 2003. Earnings management using the valuation allowance for deferred tax assets under SFAS No. 109. *Contemporary Accounting Research* 20 (3): 579–611.

Shevlin, T. 2002. Commentary on corporate tax shelters and book-tax differences. *Tax Law Review* 55 (3): 427–443.

Siegel, J. 1982. The "quality of earnings" concept—A survey. *Financial Analysts Journal* 38 (2): 60–68.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (2): 289–315.

Smith, S. 2001. Corporate tax aggressiveness and tax cushions. Working paper, Western Washington University.

Thomas, W. 2000. A test of the market's mispricing of domestic and foreign earnings. *Journal of Accounting and Economics* 28 (3): 243–267.

U.S. Department of the Treasury. 1999. *The Problem of Corporate Tax Shelters: Discussion, Analysis, and Legislative Proposals*. Washington, D.C.: Government Printing Office.

Viad Corporation. 2000. Form 10-Q. September 30. Available at http://ccbn.10kwizard.com/xml/download.php?repo=tenk&ipage=1237621&format=RTF.

Visvanathan, G. 1998. Deferred tax valuation allowances and earnings management. *Journal of Financial Statement Analysis* 3 (4): 6–15.

Walther, B. 1997. Investor sophistication and market earnings expectations. *Journal of Accounting Research* 35 (2): 157–179.

Weber, R., and J. Wheeler. 1992. Using income tax disclosures to explore significant economic transactions. *Accounting Horizons* 6 (3): 14–29.

Witt, A., and P. Behr. 2002. Enron's other strategy: Taxes; Internal papers reveal how complex deals boosted profits by $1 billion. *The Washington Post* (May 22): A-01.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (2): 357–353.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 617–652

Discretionary Accruals and Earnings Management: An Analysis of Pseudo Earnings Targets

Benjamin C. Ayers
University of Georgia

John (Xuefeng) Jiang
Michigan State University

P. Eric Yeung
University of Georgia

ABSTRACT: We investigate whether the positive associations between discretionary accrual proxies and beating earnings benchmarks hold for comparisons of groups segregated at other points in the distributions of earnings, earnings changes, and analysts-based unexpected earnings. We refer to these points as "*pseudo*" targets. Results suggest that the positive association between discretionary accruals and beating the profit benchmark extends to *pseudo* targets throughout the earnings distribution. We find similar results for the earnings change distribution. In contrast, we find few positive associations between discretionary accruals and beating *pseudo* targets derived from analysts-based unexpected earnings. We develop an additional analysis that accounts for the systematic association between discretionary accruals and earnings and earnings changes. Results suggest that the positive association between discretionary accruals and earnings intensifies around the *actual* profit benchmark (i.e., where earnings management incentives may be more pronounced). We find similar effects around the *actual* earnings increase benchmark. However, analogous patterns exist for cash flows around the profit and earnings increase benchmarks. In sum, we are unable to eliminate other plausible explanations for the associations between discretionary accruals and beating the profit and earnings increase benchmarks.

Keywords: *discretionary accruals; earnings management; earnings benchmarks.*

Data Availability: *Data are available from public sources identified in the paper.*

The authors appreciate the constructive comments of the editor, an anonymous reviewer, Steve Baginski, Linda Bamber, Dave Burgstahler, Ryan LaFond, Steve Matsunaga, John Robinson, the participants of the 2004 Southeast Summer Accounting Research Colloquium, the 2005 Financial Accounting and Reporting Section Meeting, and accounting workshops at the University of Georgia and The University of North Carolina at Charlotte. The authors also gratefully acknowledge the support of the Terry College of Business and the J. M. Tull School of Accounting.

Editor's note: This paper was accepted by Patricia Dechow.

Submitted August 2004
Accepted November 2005

I. INTRODUCTION

Prior research compares discretionary accrual proxies for firms that just beat and just miss earnings benchmarks to investigate whether firms use discretionary accruals to beat earnings benchmarks (e.g., Dechow et al. 2003; Phillips et al. 2003). One disadvantage of studies that focus solely on comparisons centered on earnings benchmarks, however, is that these comparisons may not adequately control for the possibility of a systematic association between discretionary accrual proxies and firm performance. In particular, it is difficult to interpret a positive association between discretionary accrual proxies and the likelihood of beating an earnings benchmark if they also represent firm performance.[1] This study investigates the competing explanations for the positive associations between discretionary accrual proxies and firms' propensity to beat earnings benchmarks.

In our first analysis, we use probit regressions to test whether the positive associations between discretionary accrual proxies and beating the profit, earnings increase, and analysts' forecast benchmarks hold for comparisons of groups segregated at other points in the distributions of earnings, earnings changes, and analysts-based unexpected earnings. We refer to these points as "*pseudo*" targets. We compare firm-year observations across adjacent profit "bins" and examine the association between discretionary accrual measures and the probability that a firm reports a higher profit (i.e., resides in the next higher bin). We conduct similar analyses across the distributions of earnings changes and analysts-based unexpected earnings. We use total accruals, deferred tax expense, modified Jones (1991) model abnormal accruals, and forward-looking model abnormal accruals as alternative discretionary accrual proxies. *Ex ante*, if the positive associations between discretionary accrual proxies and beating the *actual* earnings benchmarks are *solely* attributable to earnings management to beat the benchmarks, then one would expect significant positive associations between the discretionary accrual proxies and beating these *pseudo* targets *by chance* in no more than 10 percent of these comparisons where incentives to beat the *pseudo* targets should not exist.[2]

In our second analysis, we test whether the association between discretionary accrual proxies and reporting higher profits becomes more pronounced around the *actual* profit benchmarks. We conduct similar analyses for the *actual* earnings increase and analysts' forecast benchmarks. It is possible that there is an underlying significant positive association between discretionary accrual proxies and firm performance (or, alternatively that firms manage earnings throughout the earnings distribution), but the association intensifies around earnings benchmarks where earnings management incentives are stronger. Evidence of an increased association between discretionary accrual proxies and firm performance around the *actual* benchmarks would be consistent with firms using discretionary accruals to meet or beat earnings benchmarks. We test whether the positive association between discretionary accrual measures and beating each *actual* benchmark is significantly larger than (1) the *average* association between the discretionary accrual measures and beating the *pseudo* targets and (2) the association between discretionary accruals and beating *each pseudo* target.

We find significant positive associations between discretionary accrual proxies and beating the *actual* profit benchmark. However, results suggest that these positive associations extend to a significantly larger number of *pseudo* targets than expected by chance. This evidence indicates that demonstrating positive associations between a specific discretionary accrual measure and beating the profit benchmark is not sufficient to conclude that the discretionary accrual measure detects earnings management to beat the benchmark. We draw similar conclusions for the earnings increase benchmark.

---

[1] McNichols and Wilson (1988), Dechow et al. (1995), Kasznik (1999), McNichols (2000), and Dechow et al. (2003), among others, demonstrate that discretionary accruals increase with earnings.

[2] Our analyses require us to choose a probability level to quantify the number of bins that should be rejected by chance. We choose 10 percent as the probability cut-off for all statistical tests because empirical studies generally consider $p \leq .10$ to be marginally significant. Conclusions are similar when we use a more narrow probability cut-off (e.g., 5 percent).

Results from our second analysis indicate that the associations between discretionary accrual proxies and reporting higher earnings intensify around the *actual* profit benchmark. We find similar effects around the *actual* earnings increase benchmark. However, analogous patterns exist for cash flows around the profit and earnings increase benchmarks. It is possible that firms manage both cash flows and discretionary accruals around the profit and earnings increase benchmarks (Burgstahler and Dichev 1997). Alternatively, as suggested in Dechow et al. (2003), managers and employees may simply work harder around earnings benchmarks, which could explain, in part or in aggregate, the larger associations around the profit and earnings increase benchmarks. In sum, we are unable to eliminate other plausible explanations for the larger positive associations between discretionary accrual proxies and beating the profit and earnings increase benchmarks.

For the *actual analysts' forecast benchmark*, we find a significant positive association between discretionary accrual proxies and beating the benchmark. In contrast to the *pseudo* profit and earnings change analyses, we find few positive associations between discretionary accrual proxies and beating the *pseudo* unexpected earnings targets (i.e., no more than expected by chance). In addition, we find no association between changes in cash flows and beating the *actual* analysts' forecast benchmark. Thus, we are able to draw stronger conclusions regarding earnings management around the analysts' forecast benchmark.

This study makes several contributions. We demonstrate that the positive associations between discretionary accrual measures and reporting higher earnings and more positive earnings changes extend to a significantly larger number of *pseudo* targets than expected by chance. In contrast, we find few positive associations between discretionary accrual proxies and beating the *pseudo* targets derived from analysts-based unexpected earnings. We draw two conclusions from this evidence. First, the underlying associations between firm performance and discretionary accrual measures are less problematic for tests of earnings management related to analysts-based unexpected earnings. Second, demonstrating positive associations between a specific discretionary accrual measure and beating the profit and earnings increase benchmarks is not sufficient to conclude that the discretionary accrual measure detects earnings management to beat these benchmarks. Among other implications, this evidence suggests in which settings the association between discretionary accruals and performance may be problematic.

We develop an additional analysis that specifically accounts for the systematic associations between discretionary accruals and earnings and earnings changes. Results from this analysis suggest that the associations between discretionary accruals and earnings and earnings changes intensify around the *actual* profit and earnings increase benchmarks where earnings management incentives may be stronger. However, similar patterns exist for cash flows around these benchmarks. Thus, we are unable to conclude that earnings management explains the associations between discretionary accruals and beating the profit and earnings increase benchmarks.

One advantage of our analysis relative to performance-matched discretionary accruals (Kothari et al. 2005) is that it does not require specifying the exact nature of the correlation between discretionary accruals and performance—i.e., it does not require specifying a metric to match upon (e.g., ROA). Relative to performance-matching, our analysis may result in more powerful tests of earnings management around earnings benchmarks as it does not introduce "noise" in the tests via the matching process. Accordingly, this method may reduce the likelihood of falsely accepting the null hypothesis of no earnings management due to low power tests.

The remainder of this study is organized as follows. Section II presents our research method. Section III describes the sample selection, and Section IV presents descriptive statistics and results. Section V describes supplemental analyses, and Section VI concludes.

II. RESEARCH METHOD

One disadvantage of studies that focus solely on comparisons around earnings benchmarks is that they may not adequately allow for the possibility of a systematic association between discretionary accrual measures and firm performance irrespective of a firm's earnings management activities. Indeed, one of the central criticisms associated with using discretionary accrual proxies to test for earnings management is that the proxies may be capturing nondiscretionary accruals (i.e., differences in firm performance; see Dechow et al. 1995; Guay et al. 1996; Kasznik 1999; McNichols 2000). This study provides two analyses that may be useful in disentangling the competing explanations for the positive associations between discretionary accrual proxies and the propensity to meet or beat earnings benchmarks. The following paragraphs describe these analyses for the profit benchmark. Subsequent paragraphs discuss analyses for the earnings increase and analysts' forecast benchmarks.

Profit Benchmark Analyses

We test whether the positive associations between discretionary accrual measures (i.e., total accruals, deferred tax expense, modified Jones model abnormal accruals, and forward-looking model abnormal accruals) and beating the profit benchmark hold for comparisons of groups segregated based on *pseudo* profit targets.[3] Specifically, we compare firm-years across adjacent profit "bins" and examine the association between various discretionary accrual measures and the probability that a firm reports higher profits (i.e., resides in the next higher bin). Dechow et al. (1995), Kasznik (1999), and Dechow et al. (2003), among others, suggest that commonly used discretionary accrual proxies (e.g., modified Jones model abnormal accruals and forward-looking model abnormal accruals) increase in earnings. Likewise, for various firms one might anticipate a positive association between deferred tax expense and firm performance. For example, firms with depreciable assets (e.g., manufacturing firms) whose operations are not contracting are likely to maintain a constant level of or increase their depreciable assets relative to firms with poor performance. *Ceteris paribus*, new depreciable assets generate a more positive (or less negative) deferred tax balance because assets are depreciated more rapidly for tax compared to financial reporting purposes. Thus, it is plausible that there may be a positive association between deferred tax expense and firm performance.[4]

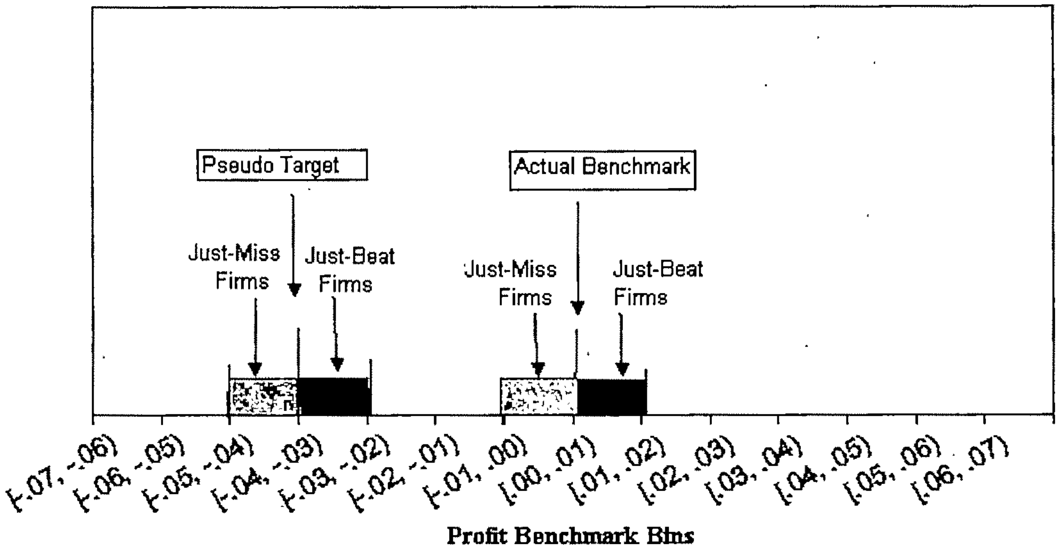
Figure 1 illustrates the comparisons across the *actual* profit benchmark and *pseudo* profit targets. To estimate the association between discretionary accrual proxies and beating

---

[3] Prior studies use total accruals (Phillips et al. 2003), deferred tax expense (Phillips et al. 2003), modified Jones model abnormal accruals (Payne and Robb 2000; Matsumoto 2002; Phillips et al. 2003), and forward-looking model accruals (Dechow et al 2003; Phillips et al. 2003) to investigate earnings management around earnings benchmarks.

[4] Other book-tax differences may lead to a less positive (or more negative) deferred tax balance (e.g., warranty expense). Thus, it is an empirical question whether there is a positive association between deferred tax expense and firm performance.

FIGURE 1
Pseudo Target and Actual Profit Benchmark Comparisons



the *actual* profit benchmark, we compare discretionary accrual proxies for firms that just-beat and just-miss the profit benchmark. In particular, we analyze the associations between discretionary accrual proxies and the firm's propensity to reside in the bin to the right of the profit benchmark. For the *pseudo* profit targets, we compare discretionary accrual proxies for firms that just-beat and just-miss the *pseudo* profit targets—i.e., we analyze the associations between discretionary accrual proxies and the firm's propensity to reside in the bin to the right of the *pseudo* profit targets.

We estimate the following probit model to test whether alternative discretionary accrual measures are positively associated with the probability that a firm meets or beats (1) the *actual* profit benchmark and (2) *pseudo* profit targets:

$$EM_{it} = \alpha + \beta_1 DisAccruals_{it} + \beta_2 CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + \varepsilon_{it} \quad (1)$$

where:

EM_{it} = 1 if $X \leq E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \leq E_{it} < X$. We define E_{it} as firm *i*'s net income in year *t* (annual Compustat data item #172) divided by the market value of equity at the end of year $t-1$ (annual Compustat data item #25 × #199), and *X* as the profit target. *X* equals 0 for the *actual* profit benchmark;[5]

$DisAccruals_{it}$ = one of the following four alternative measures of firm *i*'s discretionary accruals in year *t*:

---

[5] We use net income after extraordinary items to be consistent with previous research investigating the propensity of firms to beat earnings benchmarks (Burgstahler and Dichev 1997; Dechow et al. 2003; Phillips et al. 2003). Results are similar when we estimate the analyses using net income before extraordinary items (Compustat #123).

- *Total Accruals* = total accruals, scaled by total assets (annual Compustat #6) at the end of year $t-1$, is computed as *EBEI* − *Cash Flows*, where *EBEI* is income before extraordinary items (annual Compustat data item #123), and *Cash Flows* is cash flows from operations (computed as annual Compustat data items #308 − #124);[6]
- *Deferred Tax Expense* = deferred tax expense (annual Compustat data item #50), scaled by total assets (annual Compustat data item #6) at the end of year $t-1$;
- *Modified Jones Dis. Accruals* = abnormal accruals computed using the modified Jones model (Dechow et al. 1995; DeFond and Subramanyam 1998). We calculate *Modified Jones Dis. Accruals* as the difference between *Total Accruals* and modified Jones normal accruals. Modified Jones model normal accruals are estimated as: $Total\ Accruals_{it} = \alpha + \beta_1 (\Delta Sales_{it} - \Delta REC_{it}) + \beta_2\ PPE_{it}$, where $\Delta Sales_{it}$ equals the change in the firm's sales (annual Compustat data item #12) from year $t-1$ to year t, ΔREC_{it} equals the change in accounts receivable (annual Compustat data item #302) from year $t-1$ to year t, and *PPE* equals the firm's gross property, plant, and equipment (annual Compustat data item #7). We estimate the Modified Jones model cross-sectionally by year and industry for each two-digit SIC industry year with at least ten observations. We scale all variables by total assets at the end of year $t-1$;
- *Fwd-Look Dis. Accruals* = abnormal accruals computed using the forward-looking model (Dechow et al. 2003). We calculate *Fwd-Look Dis. Accruals* as the difference between *Total Accruals* and forward-looking normal accruals. Forward-looking model normal accruals are estimated as: $Total\ Accruals_{it} = \alpha + \beta_1 ((1 + k)\Delta Sales_{it} - \Delta REC_{it}) + \beta_2 PPE_{it} + \beta_3 Total\ Accruals_{it-1} + \beta_4 Gr\_Sales_{t+1}$, where k is the slope coefficient from a regression of ΔREC_{it} on $\Delta Sales_{it}$ (winsorized so that $0 \leq k \leq 1$) $Total\ Accruals_{it-1}$ equals total accruals from year $t-1$, scaled by total assets as of year $t-2$, and Gr_Sales_{t+1} equals the change in sales from year t to $t+1$, scaled by year t sales. We estimate the forward-looking model cross-sectionally by year and industry for each two-digit SIC industry year with at least ten observations;

$Cash\ Flows_{it}$ = cash flows from continuing operations (annual Compustat data items #308 − #124), scaled by total assets as of the end of year $t-1$;

$\Sigma_j \mathrm{Ind}_{it}$ = 1 (0) if firm i is (is not) in industry j in year t, based on two-digit SIC codes; and

ε_{it} = the error term.

Following Burgstahler and Dichev (1997), we segregate firms into "profit" bins based on the firm's year t net income deflated by end of year $t-1$ market value of equity. We design the *pseudo* targets and profit bin widths to mitigate the potential that any given firm-year observation appears in multiple just-beat or just-miss comparison groups. Each just-beat and just-miss bin has a width of .01, and each firm-year observation appears once in a just-beat group and once in a just-miss group. We limit our comparisons across *pseudo* targets to firms with deflated earnings between −.10 and .11 to ensure that we have sufficient observations per bin for our analyses and to restrict our sample to firms with less extreme earnings performance.[7] Finally, we designate *pseudo* targets in .01 increments beginning at −.09 and ending at .10. Thus, we investigate 19 *pseudo* profit targets.

---

[6] Total accruals are inherently a noisy measure of discretionary accruals. Given that prior research (e.g., Healy 1985; Phillips et al. 2003) utilizes total accruals as one measure of discretionary accruals, we include data regarding how total accruals perform across the various earnings benchmarks using both univariate and regression analyses.

Consistent with prior studies investigating discretionary accruals (e.g., Dechow et al. 1995; DeFond and Subramanyam 1998; Dechow et al. 2003; Phillips et al. 2003), we deflate each discretionary accrual measure by lag total assets. We include cash flows as a control for performance (Phillips et al. 2003) or alternatively, to control for the effect that a firm's cash flow may have on the firm's need to manage earnings via discretionary accruals. Dechow et al. (2003) estimate, based on a linear approximation of the earnings distribution, that 85 to 90 percent of firms just beat the profit benchmark without managing earnings. Probit analyses controlling for a firm's cash flows should provide more powerful tests of earnings management than univariate comparisons that do not control for the firm's need to manage earnings via discretionary accruals. Finally, we include two-digit SIC industry dummy variables in each probit regression to control for possible differences in the tendency to beat earnings targets across industries.

Ex ante, if the positive associations between discretionary accrual proxies and beating the profit benchmark are *solely* attributable to earnings management to beat the benchmark, one would expect significant positive associations between discretionary accrual proxies and beating the *pseudo* profit targets by chance in no more than 10 percent of the comparisons across *pseudo* targets (i.e., where incentives to beat the targets should not exist). This expectation is based upon using a 10 percent significance level to reject the null hypothesis. More frequent significant positive associations would be consistent with (1) a positive association between discretionary accrual proxies and firm performance (i.e., irrespective of earnings management) and/or (2) firms managing earnings upward across the distribution of earnings. Thus, it becomes difficult to distinguish between earnings management versus merely an association between discretionary accrual proxies and firm performance.

Implicit in studies analyzing discretionary accruals around earnings benchmarks is the expectation that the relationship between firm performance and discretionary accruals shifts around benchmarks. Indeed, it is possible that there is an underlying significant positive association between discretionary accrual proxies and firm performance (or, alternatively that firms manage earnings throughout the earnings distribution), but the association intensifies around earnings benchmarks where earnings management incentives are stronger. Evidence of an increased association between discretionary accrual proxies and firm performance around the *actual* profit benchmark would be consistent with firms using discretionary accruals to beat the benchmark. Accordingly, we test whether the associations between discretionary accrual proxies and reporting higher earnings are larger around the profit benchmark. We estimate the following probit model:

---

[7] Firms with deflated earnings between −.10 and .11 represent 59 percent of all firms. Similar restrictions for the earnings increase and analysts' forecast benchmark analyses (discussed later) generate samples that represent 66 percent and 73 percent of all firms, respectively. Results are similar when we expand each sample to include firms with more extreme performance (e.g., firms with deflated earnings between −.15 and .16, earnings changes between −.15 and .16, and unexpected earnings between negative 15 cents per share and positive 15 cents per share).

$$
\begin{aligned}
EM_{it} = {} & \alpha_1 + \beta_1 DisAccruals_{it} + \beta_2 CashFlows_{it} + \beta_j \Sigma_j Ind_{it} \\
& + Actual{*}(\alpha_2 + \beta_3 DisAccruals_{it} + \beta_4 CashFlows_{it} + \beta_l \Sigma_l Ind_{it}) + \varepsilon_{it} \qquad (2)
\end{aligned}
$$

where *Actual* = 1 for firm-year observations in the just-miss or just-beat bins around the *actual* profit benchmark, and 0 otherwise.

All other variables are defined in Equation (1). We estimate two different specifications of Equation (2). First, we estimate Equation (2) to determine if the associations between discretionary accrual proxies and beating the profit benchmark are significantly greater than the *average* association between the discretionary accrual proxies and beating the *pseudo* targets (i.e., we compare the association between (1) *DisAccruals* and (2) beating the *actual* profit benchmark to the respective *average* association for beating the *pseudo* targets). We include the just-beat and just-miss bins for all *pseudo* targets and the just-beat and just-miss bins for the *actual* profit benchmark in one probit regression. The coefficient for *Actual*DisAccruals* represents the incremental difference in (1) the association between the discretionary accrual proxy and beating the *actual* profit benchmark and (2) the *average* association between the discretionary accrual proxy and beating the *pseudo* targets. A significant positive coefficient for *Actual*DisAccruals* would indicate that the association between the discretionary accrual proxy and reporting higher earnings becomes more pronounced around the *actual* profit benchmark, which would be consistent with earnings management to beat the profit benchmark.

We estimate Equation (2) a second time to determine if the associations between discretionary accrual proxies and beating the profit benchmark is significantly greater than the associations between the discretionary accrual proxies and beating *each pseudo* target (i.e., we compare the associations between (1) *DiscAccruals* and (2) beating the *actual* profit benchmark to the respective associations for *each pseudo* target). We estimate Equation (2) once for each specific *pseudo* profit target. In these analyses, there are four bins included in each analysis. We include the just-miss and just-beat bins around the *actual* profit benchmark and the just-miss and just-beat bins around the *pseudo* profit target. Similar to above, a significant positive coefficient for *Actual*DisAccruals* would indicate that the association between the discretionary accrual proxy and reporting higher earnings becomes more pronounced around the *actual* benchmarks, which would be consistent with earnings management to beat the profit benchmark.

Earnings Increase Benchmark Analyses

We make three changes to Equations (1) and (2) to investigate the associations between discretionary accruals and beating the *actual* earnings increase benchmark and *pseudo* earnings change targets. First, we segregate firms into "earnings change" bins based on the firm's year t earnings change deflated by end of year $t-2$ market value of equity. Similar to the profit benchmark analyses, each just-beat and just-miss bin has a width of .01, and each firm-year observation appears once in a just-beat group and once in a just-miss group. We investigate 19 *pseudo* targets consisting of firms with deflated earnings changes between $-.10$ and .11. Second, we redefine EM_{it} to equal 1 if $X \leq \Delta E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \leq \Delta E_{it} < X$. We define ΔE_{it} as the change in firm i's net income from year $t-1$ to t divided by the market value of equity at the end of year $t-2$, and X as the earnings change target. X equals 0 for the *actual* earnings increase benchmark. Third, we replace *CashFlows* in Equations (1) and (2) with Δ*CashFlows* to control for performance (Phillips et al. 2003). We define Δ*CashFlows* as the change in firm i's cash flows from continuing operations (annual Compustat data items #308 − #124) from year $t-1$ to t, scaled by total assets at the end of year $t-1$.

Analysts' Forecast Benchmark Analyses

We make three modifications to Equations (1) and (2) for the analysts' forecast benchmark analyses. First, we assign firms to "analysts-based unexpected earnings" bins based on the firm's unexpected earnings per share (in cents). We define UE_{it} (unexpected earnings) as firm i's year t actual earnings per share minus the single-most recent analyst forecast provided prior to the earnings announcement, both available from the Unadjusted I/B/E/S Detail History file. Consistent with prior research (e.g., Degeorge et al. 1999; Payne and Thomas 2003; Phillips et al. 2003), we round UE_{it} to the nearest penny. Each just-beat and just-miss bin has a width of .01, and each firm-year observation appears once in a just-beat group and once in a just-miss group. We investigate 19 *pseudo* targets in these analyses consisting of firms with unexpected earnings between negative ten cents per share and positive ten cents per share. Second, we redefine EM_{it} to equal 1 if UE_{it} equals X cents per share, and 0 if UE_{it} equals X less one cent per share. We define X as the unexpected earnings target, and X equals 0 for the *actual* analysts' forecast benchmark.[8] Third, we replace *CashFlows* in Equations (1) and (2) with $\Delta CashFlows$ in these analyses to control for performance (Phillips et al. 2003).

We use the single-most recent analyst forecast issued prior to the earnings announcement from the Unadjusted I/B/E/S Detail History file as our forecast benchmark. O'Brien (1988) and Brown (1991) suggest that the single-most recent analyst forecast is more accurate in predicting actual earnings than the consensus mean forecast. Likewise, Brown and Kim (1991) find that the single-most recent analyst forecast more accurately reflects the market's earnings expectation than the consensus mean forecast. Assuming that firms intend to meet or beat market expectations, using a more current forecast proxy should provide a more powerful test of whether firms use discretionary accruals to meet or beat analyst forecasts. This may be especially important since recent research suggests that firms "walk down" analyst forecasts during the sample period (Richardson et al. 2004).

III. SAMPLE SELECTION

We consider four discretionary accrual measures in our tests, including total accruals, deferred tax expense, modified Jones model abnormal accruals, and forward-looking model abnormal accruals. To facilitate a consistent measure of deferred tax expense across our sample period, we restrict the sample to post-SFAS No. 109 deferred tax expense measures (i.e., firm-years from 1994–2002). We exclude firms not incorporated in the U.S., mutual funds (SIC code 6726), trusts (SIC code 6792), REITS (SIC code 6798), limited partnerships (SIC code 6799), and other flow-through entities (SIC code 6795) because these firms are either primarily subject to taxation outside the U.S. or do not account for income tax expense. We exclude utilities (SIC codes 4900–4999) and financial institutions (SIC codes 6000–6099) as regulated firms likely face different earnings management incentives than nonregulated firms. To mitigate the effect of extreme observations, we delete firm-years with a deferred tax expense in the top or bottom one percentile in each comparison or with a scaled total accrual, modified Jones model abnormal accruals, or forward-looking model abnormal accruals greater than 100 percent (in absolute value) of lagged total assets (DeFond and Subramanyam 1998; Dechow et al. 2003; Phillips et al. 2003). Finally, each firm-year observation must have sufficient information available to calculate the requisite variables in the probit analysis.

---

[8] Results are similar when we (1) designate firms with *UE* equal to one cent per share as the just-beat group and eliminate sample firms with *UE* equal to zero cents per share or (2) designate firms with *UE* equal to zero or one cent per share as the just-beat group.

IV. DESCRIPTIVE STATISTICS AND RESULTS

Profit Benchmark Analyses

Table 1, Panel A presents the sample size for each profit bin and the mean values of each discretionary accrual measure. In each row, we test whether the mean values of the discretionary accrual measures for that particular profit bin is significantly higher (i.e., more positive or less negative) than the mean values of the discretionary accrual measures in the preceding row. We denote significant differences ($p \leq .10$, one-tailed test) with bold text and identify the "*actual*" just-beat and just-miss groups with "boxed" text. To provide additional evidence regarding the potential positive correlation between discretionary accrual proxies and firm performance, the last row of Panel A reports the correlation between the mean values of discretionary accrual measures for each bin and the bin number, which ranges from 1 (the most negative bin) to 21 (the most positive bin). Table 1, Panel B presents the Pearson correlations for each variable (i.e., not the mean values per bin).

The profit benchmark sample includes 22,903 firm-year observations when *Total Accruals*, *Deferred Tax Expense*, or *Modified Jones Dis. Accruals* proxies for discretionary accruals and 19,045 firm-year observations when *Fwd-Look Dis. Accruals* proxies for discretionary accruals. The *Fwd-Look Dis. Accruals* sample is smaller because estimating *Fwd-Look Dis. Accruals* requires two additional years of data (i.e., to estimate $Total\ Accruals_{it-1}$ and Gr_Sales_{t+1}). 1,180 observations reside in the *actual* just-beat bin (i.e., $.00 \leq E_{it} < .01$) and 683 reside in the *actual* just-miss bin (i.e., $-.01 \leq E_{it} < .00$) when *Total Accruals*, *Deferred Tax Expense*, or *Modified Jones Dis. Accruals* proxies for discretionary accruals. 963 observations reside in the *actual* just-beat bin and 548 reside in the *actual* just-miss bin when *Fwd-Look Dis. Accruals* proxies for discretionary accruals. Similar to prior research, when we graph the frequency of bin observations for the profit benchmark sample, we observe a "kink" in the distributions to the left of the zero profit bin.

Panel A of Table 1 indicates that *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals* are significantly larger (i.e., more positive or less negative at $p \leq .10$, one-tailed test) in the $.00 \leq E_{it} < .01$ bin (i.e., the just-beat bin) relative to the $-.01 \leq E_{it} < .00$ bin (i.e., the just-miss bin).[9] Nonetheless, similar results hold for five of the 19 other comparisons for *Total Accruals*, seven of the 19 other comparisons for *Deferred Tax Expense*, and four of the other 19 comparisons for *Fwd-Look Dis. Accruals* ($p \leq .10$, one-tailed test). These univariate comparisons suggest that the positive associations between discretionary accrual measures and earnings extend to other earnings bins not centered on the zero profit benchmark. We also note, as might be expected, that *CashFlows* increases in earnings. In ten of 20 comparisons, the mean value of *CashFlows* is statistically larger (more positive or less negative at $p \leq .10$) than its value in the preceding bin.

---

[9] Prior research (Dechow et al. 2000, 2003; Phillips et al. 2003) report similar results for univariate comparisons of *Total Accruals*, *Deferred Tax Expense*, and *Modified Jones Dis. Accruals* for just-beat and just-miss firms around the *actual* profit benchmark. For *Fwd-Look Dis. Accruals*, both Dechow et al. (2003) and we report a significantly larger amount for just-beat firms ($p = .07$ and $p = .01$, respectively, one-tailed test), while Phillips et al. (2003) report no significant difference. When we limit our sample to the period investigated in Dechow et al. (2003), we find a significantly larger *Fwd-Look Dis. Accruals* for just-beat firms ($p = .06$, one-tailed test).

TABLE 1
Profit Benchmark Analyses

Panel A: Descriptive Statistics

E_{it}	n[a]	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*CashFlows*
$-.10 \le E_{it} < -.09$	421/319	−0.1102	−0.0056	−0.0280	−0.0282	−0.0147
$-.09 \le E_{it} < -.08$	437/344	−0.0987	−0.0065	−0.0270	−0.0230	−0.0109
$-.08 \le E_{it} < -.07$	457/343	−0.0984	−0.0068	−0.0290	−0.0123	−0.0187
$-.07 \le E_{it} < -.06$	505/377	−0.0991	−0.0064	−0.0288	−0.0251	−0.0094
$-.06 \le E_{it} < -.05$	551/438	**−0.0882**	−0.0069	−0.0237	**−0.0087**	**0.0003**
$-.05 \le E_{it} < -.04$	588/458	−0.0802	−0.0060	−0.0165	−0.0050	0.0027
$-.04 \le E_{it} < -.03$	648/492	−0.0735	−0.0061	−0.0155	0.0005	**−0.0090**
$-.03 \le E_{it} < -.02$	667/542	−0.0747	−0.0065	**0.0084**	−0.0059	**0.0085**
$-.02 \le E_{it} < -.01$	727/577	**−0.0625**	−0.0072	0.0064	**0.0041**	**0.0092**
$-.01 \le E_{it} < .00$	683/548	−0.0598	**−0.0050**	**0.0210**	0.0050	**0.0083**
$.00 \le E_{it} < .01$	1180/963	**−0.0355**	**−0.0032**	0.0301	**0.0282**	**0.0023**
$.01 \le E_{it} < .02$	1301/1049	−0.0396	**−0.0010**	0.0278	0.0273	**0.0114**
$.02 \le E_{it} < .03$	1535/1255	−0.0375	**0.0000**	0.0218	0.0173	**0.0232**
$.03 \le E_{it} < .04$	1737/1432	**−0.0330**	**0.0018**	**0.0316**	0.0194	**0.0192**
$.04 \le E_{it} < .05$	1982/1678	−0.0340	**0.0025**	0.0296	0.0195	**0.0267**
$.05 \le E_{it} < .06$	2133/1866	**−0.0297**	**0.0032**	0.0273	**0.0234**	0.0172
$.06 \le E_{it} < .07$	2097/1819	−0.0274	0.0032	0.0273	0.0197	0.0173
$.07 \le E_{it} < .08$	1776/1540	−0.0295	0.0031	0.0218	0.0200	0.0227
$.08 \le E_{it} < .09$	1441/1241	−0.0275	0.0027	0.0247	0.0190	0.0241
$.09 \le E_{it} < .10$	1118/958	**−0.0193**	**0.0034**	**0.0365**	**0.0298**	0.0194
$.10 \le E_{it} < .11$	919/806	−0.0207	0.0021	0.0302	0.0225	0.0227
Pearson Correlation between Mean Discretionary Accrual Measures and Bin Number		0.97	0.92	0.90	0.89	0.91
(p-value, two-tailed)		(.00)	(.00)	(.00)	(.00)	(.00)

(*continued on next page*)

TABLE 1 (Continued)

Panel B: Pearson Correlations

	Total Accruals	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*Cash Flows*
E	**0.2205**	**0.2426**	**0.0761**	**0.1088**	**0.4428**
Total Accruals		**0.0354**	**0.4382**	**0.6264**	**−0.3884**
Deferred Tax Expense			**0.0361**	**0.0507**	**0.0599**
Modified Jones Dis. Accruals				**0.4641**	**−0.2023**
Fwd-Look Dis. Accruals					**−0.2794**

Panel C: Probit Analyses across Adjacent Profit Groups

$$EM_{it} = \alpha + \beta_1 DisAccruals_{it} + \beta_2 CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + \varepsilon_{it} \quad (1)$$

Just-Beat Sample	Just-Miss Sample	n^a	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*CashFlows*
$-.09 \leq E_{it} < -.08$	$-.10 \leq E_{it} < -.09$	852/654	0.370^c (0.14)	0.949^c (0.36)	−0.061^c (0.61)	0.110^c (0.38)	−0.163^c (0.75)
$-.08 \leq E_{it} < -.07$	$-.09 \leq E_{it} < -.08$	887/679	0.018^c (0.48)	−0.132^c (0.52)	0.121 (0.29)	**0.456^c** **(0.10)**	0.064^c (0.39)
$-.07 \leq E_{it} < -.06$	$-.08 \leq E_{it} < -.07$	952/707	−0.087^c (0.60)	1.662^c (0.23)	0.118 (0.29)	−0.452^c (0.87)	0.293^c (0.11)
$-.06 \leq E_{it} < -.05$	$-.07 \leq E_{it} < -.06$	1036/803	**0.634^c** **(0.04)**	−2.537^c (0.83)	0.109 (0.31)	**0.531^c** **(0.07)**	**0.577^c** **(0.01)**
$-.05 \leq E_{it} < -.04$	$-.06 \leq E_{it} < -.05$	1117/878	**0.589^c** **(0.05)**	2.775^d (0.14)	0.160 (0.20)	0.132^c (0.34)	0.097^c (0.32)
$-.04 \leq E_{it} < -.03$	$-.05 \leq E_{it} < -.04$	1226/941	**0.466^c** **(0.09)**	−0.645^c (0.61)	0.109 (0.28)	0.267^c (0.23)	**0.294^c** **(0.07)**
$-.03 \leq E_{it} < -.02$	$-.04 \leq E_{it} < -.03$	1294/1013	0.329^c (0.19)	0.118^c (0.48)	**0.430** **(0.01)**	0.165^c (0.33)	**1.523** **(0.00)**
$-.02 \leq E_{it} < -.01$	$-.03 \leq E_{it} < -.02$	1380/1101	**1.637^c** **(0.00)**	−1.915^c (0.79)	0.096 (0.31)	**1.038** **(0.00)**	**1.100** **(0.00)**
$-.01 \leq E_{it} < .00$	$-.02 \leq E_{it} < -.01$	1400/1107	**2.484^c** **(0.00)**	**7.654** **(0.00)**	**0.619** **(0.00)**	**0.777^c** **(0.03)**	**1.611** **(0.00)**

(continued on next page)

TABLE 1 (Continued)

$.00 \le E_{it} < .01$	$-.01 \le E_{it} < .00$	1848/1494	**41.054** **(0.00)**	**8.510** **(0.00)**	**0.431** **(0.01)**	**1.686** **(0.00)**	**1.536** **(0.00)**
$.01 < E_{it} < .02$	$.00 \le E_{it} < .01$	2471/1996	**9.652**^{c} **(0.00)**	**8.538** **(0.00)**	**0.462** **(0.00)**	**1.310** **(0.00)**	**2.359** **(0.00)**
$.02 \le E_{it} < .03$	$.01 \le E_{it} < .02$	2803/2266	**6.514**^{c} **(0.00)**	**3.039**^{c} **(0.05)**	**0.250** **(0.06)**	**0.425**^{c} **(0.09)**	**1.833** **(0.00)**
$.03 \le E_{it} < .04$	$.02 \le E_{it} < .03$	3245/2667	**2.788**^{c} **(0.00)**	**5.144** **(0.00)**	**0.477** **(0.00)**	**0.954**^{d} **(0.00)**	**1.444** **(0.00)**
$.04 \le E_{it} < .05$	$.03 \le E_{it} < .04$	3702/3089	**1.932**^{c} **(0.00)**	**2.173**^{c} **(0.09)**	**0.241** **(0.06)**	**0.976**^{d} **(0.00)**	**1.534** **(0.00)**
$.05 \le E_{it} < .06$	$.04 \le E_{it} < .05$	4098/3512	**0.502**^{c} **(0.10)**	**2.385**^{c} **(0.07)**	−0.124^{c} (0.78)	**0.540**^{c} **(0.04)**	−0.129^{c} (0.69)
$.06 \le E_{it} < .07$	$.05 \le E_{it} < .06$	4210/3655	**0.890**^{c} **(0.01)**	−0.083^{c} (0.52)	0.172 (0.18)	−0.175^{c} (0.71)	0.314^{c} (0.13)
$.07 \le E_{it} < .08$	$.06 \le E_{it} < .07$	3841/3328	−0.779^{c} (0.97)	0.656^{c} (0.35)	−0.161^{c} (0.80)	0.405^{c} (0.11)	0.166^{c} (0.29)
$.08 \le E_{it} < .09$	$.07 \le E_{it} < .08$	3198/2757	**0.977**^{c} **(0.02)**	−0.757^{c} (0.67)	0.238 (0.13)	0.376^{c} (0.14)	**0.505**^{c} **(0.05)**
$.09 \le E_{it} < .10$	$.08 \le E_{it} < .09$	2549/2183	−0.245^{c} (0.69)	2.037^{c} (0.14)	**0.298** **(0.10)**	**0.757**^{c} **(0.03)**	−0.145^{c} (0.66)
$.10 \le E_{it} < .11$	$.09 \le E_{it} < .10$	2027/1758	0.368^{c} (0.24)	−5.506^{c} (0.99)	−0.192^{c} (0.77)	−0.322^{c} (0.78)	0.193^{c} (0.31)
Number of Significant Positive Coefficients for Pseudo Targets^{b}			12 (0.00)	6 (0.01)	7 (0.00)	10 (0.00)	10 (0.00)

(*continued on next page*)

TABLE 1 (Continued)

Panel D: Pooled Probit Analyses Comparing the Actual Profit Benchmark to All other Pseudo Targets

$$EM_{it} = \alpha_1 + \beta_1 DisAccruals_{it} + \beta_2 CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + Actual{*}(\alpha_2 + \beta_2 DisAccruals_{it} + \beta_4 CashFlows_{it} + \beta_l \Sigma_l Ind_{it}) + \varepsilon_{it} \quad (2)$$

	***DisAccruals* is defined as:**			
Independent Variables	***Total Accruals***	***Deferred Tax Expense***	***Modified Jones Dis. Accruals***	***Fwd-Look Dis. Accruals***
DisAccruals	0.514	1.395	0.127	0.323
	(0.00)	(0.00)	(0.00)	(0.00)
CashFlows	0.503	0.345	0.374	0.487
	(0.00)	(0.00)	(0.00)	(0.00)
*Actual*DisAccruals*	40.540	7.115	0.304	1.363
	(0.00)	(0.00)	(0.06)	(0.00)
*Actual*CashFlows*	38.960	0.345	0.553	1.049
	(0.00)	(0.13)	(0.05)	(0.01)
n	44,136	44,136	44,136	36,588

We construct probability levels for all statistical tests in Panels A, C, and D using one-tailed tests. Panels A and C highlight the "*actual*" just-beat and just-miss groups for profit benchmark with "boxed" text.

Bold text in Panel A indicates that the mean discretionary accruals value is significantly larger than the mean value of the preceding group ($p \leq .10$, one-tailed test). **Bold** text in Panel B highlights the significant Pearson correlation (($p \leq .10$, two-tailed test). **Bold** text in Panel C highlights significant coefficient for discretionary accruals measures ($p \leq .10$, one-tailed test).

[a] The number to the left of the slash "/" identifies the sample size for the analyses when *Total Accruals*, *Deferred Tax Expense*, and *Modified Jones Dis. Accruals* proxy for discretionary accruals. The number to the right of the slash "/" is the sample size for the analyses when *Fwd-Look Dis. Accruals* proxy for discretionary accruals.

[b] In the last two rows of Panel C we report (1) the number of significant positive coefficients (*k*) for each discretionary accrual measures ($p \leq .10$, one-tailed test) and (2) the probability of "drawing" *k* or more significant positive coefficients assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance.

[c] or [d] in Panel C indicate that the coefficient from the probit regression using the sample restricted to just-beat and just-miss firms around the *actual* benchmark is significantly larger than the estimated coefficient for this specific *pseudo* target at $p \leq .05$ (one-tailed test) or $p \leq .10$ (one-tailed test), respectively.

(*continued on next page*)

TABLE 1 (Continued)

Variable Definitions:

E_{it} = net income (annual Compustat data item #172) in year t scaled by the market value of equity at the end of year $t-1$ (annual Compustat data item #25 × #199);

Total Accruals = total accruals, scaled by total assets (annual Compustat #6) at the end of year $t-1$, is computed as *EBEI* − *Cash Flows*, where *EBEI* is income before extraordinary items (annual Compustat data item #123), and *Cash Flows* is cash flows from operations (computed as annual Compustat data items #308 − #124);

Deferred Tax Expense = deferred tax expense (annual Compustat data item #50), scaled by total assets (annual Compustat data item #6) at the end of year $t-1$;

Modified Jones Dis. Accruals = abnormal accruals computed using the modified Jones model (Dechow et al. 1995; DeFond and Subramanyam 1998). We calculate *Modified Jones Dis. Accruals* as the difference between *Total Accruals* and modified Jones normal accruals. Modified Jones model normal accruals are estimated as: $Total\ Accruals_{it} = \alpha + \beta_1 (\Delta Sales_{it} - \Delta REC_{it}) + \beta_2 PPE_{it}$, where $\Delta Sales_{it}$ equals the change in the firm's sales (annual Compustat data item #12) from year $t-1$ to year t, ΔREC_{it} equals the change in accounts receivable (annual Compustat data item #302) from year $t-1$ to year t, and *PPE* equals the firm's gross property, plant, and equipment (annual Compustat data item #7). We estimate the Modified Jones model cross-sectionally by year and industry years for each two-digit SIC industry year with at least ten observations. We scale all variables by total assets at the end of year $t-1$;

Fwd-Look Dis. Accruals = abnormal accruals computed using the forward-looking model (Dechow et al. 2003). We calculate *Fwd-Look Dis. Accruals* as the difference between *Total Accruals* and forward-looking normal accruals. Forward-looking model normal accruals are estimated as: $Total\ Accruals_{it} = \alpha + \beta_1 ((1 + k)\Delta Sales_{it} - \Delta REC_{it}) + \beta_2 PPE_{it} + \beta_3 Total\ Accruals_{it-1} + \beta_4 Gr\_Sales_{t+1}$, where k is the slope coefficient from a regression of ΔREC_{it} on $\Delta Sales_{it}$, $Total\ Accruals_{it-1}$ equals total accruals from year $t-1$, scaled by total assets as of year $t-2$, and Gr_Sales_{t+1} equals the change in sales from year t to $t+1$, scaled by year t sales. We estimate the forward-looking model cross-sectionally by year and industry for each two-digit SIC industry year with at least ten observations;

CashFlows = cash flows from operations (annual Compustat data items #308 − #124), scaled by total assets as of the end of year $t-1$;

EM_{it} = 1 if $X \leq E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \leq E_{it} < X$. X is the profit target. X equals 0 for the *actual* profit benchmark;

$\Sigma_j Ind_{it}$ = 1 (0) if firm i is (is not) in industry j in year t, based on two-digit SIC codes; and

Actual = 1 for firm-year observations in the just-miss or just-beat bins around the *actual* benchmarks, and 0 otherwise.

The last row of Table 1, Panel A indicates that the mean values of the discretionary accrual proxies are highly positively correlated with the earnings bin number. The Pearson correlation coefficients for *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are .97, .92, .90, and .89, respectively, and each is statistically significant ($p \leq .00$). Panel B presents the Pearson correlation coefficients for each variable. *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are each positively correlated with earnings (E_{it}) ($p \leq .10$, two-tailed test). In sum, univariate analyses suggest that the associations between discretionary accrual proxies and reporting higher earnings may not be unique to comparisons centered on the profit benchmark.

Comparisons across Pseudo Targets

Table 1, Panel C reports the results of the *pseudo* target and *actual* benchmark analyses.[10] The fourth through seventh columns present the coefficients for the alternative discretionary accrual measures. In the last column we present the coefficients for *Cash Flows* from the regression including *Fwd-Look Dis. Accruals* as the discretionary accrual proxy. The pattern of significant coefficients for *Cash Flows* is similar when we use *Total Accruals*, *Deferred Tax Expense*, or *Modified Jones Dis. Accruals* as the discretionary accrual proxy. Each row represents the results for estimating Equation (1) for a different just-beat and just-miss sample. We highlight positive and significant coefficients ($p \leq .10$, one-tailed test) with bold text and identify the *actual* just-beat and just-miss groups with boxed text. In the last two rows we tally the number of significant positive coefficients (*k*) for the *pseudo* target regressions for each discretionary accrual measure ($p \leq .10$, one-tailed test) and present the probability of "drawing" *k* or more significant positive coefficients assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance.

We find in Panel C of Table 1 that *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are each positively associated with the probability of beating the *actual* profit benchmark ($p \leq .01$, one-tailed test). Results suggest, however, that the significant positive associations between these discretionary accrual proxies and firm profits exist for more than the expected number of comparisons of adjacent bins segregated based on *pseudo* profit targets. Of the 19 comparisons across adjacent bins segregated based on *pseudo profit* targets, we find a significant positive association ($p \leq .10$, one-tailed test) between beating the *pseudo* target and (1) *Total Accruals* in 12 comparisons (63 percent), (2) *Deferred Tax Expense* in six comparisons (32 percent), (3) *Modified Jones Dis. Accruals* in seven comparisons (37 percent), and (4) *Fwd-Look Dis. Accruals* in ten comparisons (52 percent). Assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance, the frequency of positive significant associations for each discretionary accrual measure is significantly higher than expected ($p \leq .01$, one tailed test). Accordingly, it is difficult to distinguish whether the associations between *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* and beating the *actual* profit benchmark are attributable to earnings management or the associations between these discretionary accrual proxies and firm performance. Finally, as might be expected, Panel C indicates that *CashFlows* is positively related to beating the *actual* profit benchmark and is also positively associated with beating ten of the 19 *pseudo* profit targets ($p \leq .10$, one tailed test).

---

[10] Our probit regression results comparing just-beat and just-miss firms around the *actual* profit benchmark are similar to the probit results reported in Phillips et al. (2003).

Tests for Shifts in the Association between Discretionary Accrual Measures and Firm Performance around the Profit Benchmark

We estimate Equation (2) to test whether the associations between discretionary accrual proxies and firm performance increase around the *actual* profit benchmark (i.e., where earnings management incentives may be stronger). In Table 1, Panel C, we identify with superscripts c and d those regression coefficients from the *pseudo* target analyses that are significantly *smaller* ($p \leq .05$ and $p \leq .10$, one-tailed test, respectively) than the estimated coefficient from the *actual* profit benchmark (i.e., significantly smaller than the coefficients highlighted in boxed text).

In Table 1, Panel D, we report the results when we compare the associations between discretionary accrual proxies and beating the *actual* profit benchmark to the *average* association between the discretionary accrual proxies and beating the *pseudo* targets. Comparing the coefficients for *DisAccruals* estimated for the actual benchmark and pseudo targets via *Actual*DisAccruals* in Equation (2) is similar to comparing the marginal effects of the discretionary accrual measures for the actual benchmark and pseudo targets. In particular, comparisons of the marginal effects for *DisAccruals* for the actual benchmarks and pseudo targets generate the same sign and statistical significance as *Actual*DisAccruals*. Likewise, the ratio of one coefficient (e.g., *DisAccruals*) to another (e.g., *CashFlows*) is similar in magnitude to the ratio of their marginal effects, which allows one to compare the relative marginal effects of each variable.

Consistent with the relationship between discretionary accrual proxies and firm performance increasing around the profit benchmark, Table 1, Panel D reports that the coefficients for *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* around the *actual* profit benchmark are significantly larger than the *average* association between the respective discretionary accrual measures and beating the *pseudo* profit targets. In particular, the coefficients for *Actual*Total Accruals*, *Actual*Deferred Tax Expense*, *Actual*Modified Jones Dis. Accruals*, and *Actual*Fwd-Look Dis. Accruals* are positive and significant ($p = .00$, .00, .06, and .00, respectively, one-tailed test).

In addition, Table 1, Panel C reports that the coefficients for *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals* estimated around the *actual* profit benchmark are significantly larger than the majority of the estimated coefficients from the *pseudo* profit target comparisons. In particular, the estimated coefficients around the profit benchmark are significantly larger than 19 of the 19 coefficients for *Total Accruals* for the *pseudo* targets, 16 of the 19 coefficients for *Deferred Tax Expense* for the *pseudo* targets, and 17 of the 19 coefficients for *Fwd-Look Dis. Accruals* for the *pseudo* targets ($p < .10$, one-tailed test). These frequencies are significantly higher than expected ($p = .00$, one-tailed test) assuming a binomial distribution and 10 percent likelihood that the *actual* benchmark coefficient is significantly larger than a *pseudo* target coefficient by chance. We find that the estimated coefficient for *Modified Jones Dis. Accruals* for the profit benchmark is significantly larger ($p < .10$, one-tailed test) than four of the 19 coefficients for *Modified Jones Dis. Accruals* for the *pseudo* targets.

In sum, results suggest that the positive associations between discretionary accrual proxies and firm profits intensify around the *actual* profit benchmark. These findings are consistent with the earnings management hypothesis—i.e., the association between discretionary accrual proxies and firm performance shifts in those settings where firms have greater incentives to manage earnings. However, the associations between *CashFlows* and firm profits generally also increase around the *actual* profit benchmark. Specifically, Table

1, Panel D reports that the association between *CashFlows* and beating the *actual* profit benchmark is significantly larger ($p \leq .05$, one-tailed test) than the *average* association between *CashFlows* and beating the *pseudo* profit targets in three of the four specifications of Equation (2). Likewise, Table 1, Panel C indicates that the coefficient for *CashFlows* from the regression estimated around the *actual* profit benchmark is significantly larger than most of the coefficients for *CashFlows* for the *pseudo* earnings targets.

There are at least three possible explanations for the increased association between *CashFlows* and earnings around the *actual* profit benchmark. First, it is possible that firms manage both cash flows and discretionary accruals around the profit benchmark (Burgstahler and Dichev 1997). Indeed, evidence reported by Roychowdhury (2003) suggests that firms manage cash flows from operations to avoid reporting losses. Second, it is possible that cash flows from operations better represent firm performance around the profit benchmark. For example, the association between cash flows from operations and net income may increase around the profit benchmark if firms are less likely to report extraordinary items around the profit benchmark. We examined the distribution of extraordinary items for our sample and note that extraordinary items are more pronounced at the tails of the earnings distributions. In sensitivity analysis, we re-estimated our probit regressions after limiting the sample to firms that do not report extraordinary items. Results are similar to those reported in Table 1, Panels C and D for our discretionary accrual measures. In addition, the coefficients for *CashFlows* remain more pronounced around the profit benchmark. Third, as suggested in Dechow et al. (2003), it is possible that managers and employees may simply work harder around earnings benchmarks. If this conjecture is descriptive, then it could, in part or in aggregate, explain the unique association between discretionary accruals and beating the profit benchmark. Accordingly, we are unable to conclude with confidence that earnings management explains this association.

Earnings Increase Benchmark Analyses

Table 2, Panels A and B, report univariate analyses for the earnings increase benchmark. In each row of Panel A, we test whether the mean values of the discretionary accrual measures for that particular earnings change bin is significantly higher (i.e., more positive or less negative) than the mean values of the discretionary accrual measures in the preceding row. Panel A indicates that both *Total Accruals* and *Deferred Tax Expense* are significantly larger (i.e., more positive or less negative at $p \leq .10$, one-tailed test) in the $.00 \leq \Delta E_{it} < .01$ bin (i.e., the *actual* just-beat bin) relative to the $-.01 \leq \Delta E_{it} < .00$ bin (i.e., the *actual* just-miss bin).[11] However, these same relations hold for seven of the 19 other comparisons for *Total Accruals* and three of the 19 other comparisons for *Deferred Tax Expense* ($p \leq .10$, one-tailed test).

The last row of Table 2, Panel A indicates that the mean values of the discretionary accruals are highly positively correlated with the earnings change bin number. The Pearson correlation coefficients for *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are .91, .83, .85, and .94, respectively, and each is statistically significant ($p \leq .00$). Panel B presents the Pearson correlation coefficients for each variable. *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are each positively correlated with earnings changes (ΔE_{it}) ($p \leq .10$, two-tailed test). In sum, univariate analyses suggest that the associations between

---

[11] Phillips et al. (2003) report similar results for univariate comparisons of *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* for just-beat and just-miss firms around the *actual* earnings increase benchmark.

TABLE 2
Earnings Increase Benchmark Analyses

Panel A: Descriptive Statistics

ΔE_{it}	n[a]	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*ΔCash Flows*
$-.10 \leq \Delta E_{it} < -.09$	366/307	−0.1032	−0.0058	−0.0383	−0.0344	−0.0531
$-.09 < \Delta E_{it} < -.08$	449/372	−0.1064	−0.0053	−0.0434	−0.0396	−0.0260
$-.08 \leq \Delta E_{it} < -.07$	504/399	−0.0975	−0.0040	**−0.0253**	**−0.0227**	−0.0328
$-.07 \leq \Delta E_{it} < -.06$	573/467	−0.0933	−0.0030	−0.0392	−0.0305	−0.0352
$-.06 \leq \Delta E_{it} < -.05$	710/603	**−0.0815**	−0.0037	**−0.0222**	**−0.0194**	**−0.0237**
$-.05 \leq \Delta E_{it} < -.04$	820/658	−0.0840	**−0.0019**	−0.0198	−0.0191	−0.0186
$-.04 \leq \Delta E_{it} < -.03$	1029/826	**−0.0730**	−0.0021	−0.0117	**−0.0110**	−0.0153
$-.03 \leq \Delta E_{it} < -.02$	1279/1036	**−0.0650**	**−0.0005**	−0.0031	**−0.0033**	−0.0185
$-.02 \leq \Delta E_{it} < -.01$	1585/1312	**−0.0591**	−0.0004	0.0030	−0.0001	**−0.0057**
$-.01 \leq \Delta E_{it} < .00$	1995/1636	**−0.0531**	**0.0009**	**0.0183**	**0.0129**	**0.0016**
$.00 \leq \Delta E_{it} < .01$	2942/2471	**−0.0482**	**0.0020**	0.0110	0.0112	**0.0210**
$.01 \leq \Delta E_{it} < .02$	2618/2220	**−0.0387**	0.0023	**0.0175**	**0.0152**	**0.0265**
$.02 \leq \Delta E_{it} < .03$	1878/1586	−0.0448	0.0021	0.0147	0.0112	**0.0331**
$.03 \leq \Delta E_{it} < .04$	1410/1181	−0.0442	0.0017	0.0122	**0.0168**	**0.0478**
$.04 \leq \Delta E_{it} < .05$	1032/887	−0.0470	0.0020	0.0098	0.0131	0.0483
$.05 \leq \Delta E_{it} < .06$	848/706	−0.0431	0.0017	0.0128	**0.0214**	0.0380
$.06 \leq \Delta E_{it} < .07$	718/606	**−0.0341**	0.0020	0.0113	0.0254	**0.0496**
$.07 \leq \Delta E_{it} < .08$	565/465	−0.0430	0.0015	0.0186	0.0145	**0.0627**
$.08 \leq \Delta E_{it} < .09$	511/432	−0.0403	0.0010	0.0110	0.0248	0.0590
$.09 \leq \Delta E_{it} < .10$	401/330	−0.0348	0.0014	0.0195	0.0271	0.0517
$.10 \leq \Delta E_{it} < .11$	387/319	−0.0477	0.0002	0.0052	0.0230	0.0651
Pearson Correlation between Mean Discretionary Accrual Measures and Bin Number		0.91	0.83	0.85	0.94	0.97
(p-value, two-tailed)		(.00)	(.00)	(.00)	(.00)	(.00)

(continued on next page)

TABLE 2 (Continued)

Panel B: Pearson Correlations

	Total Accruals	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*Cash Flows*
ΔE	**0.1590**	**0.1267**	**0.0624**	**0.1374**	**0.1824**
Total Accruals		0.0010	**0.4565**	**0.6469**	**−0.3129**
Deferred Tax Expense			**0.0278**	**0.0356**	**0.0100**
Modified Jones Dis. Accruals				**0.4675**	**−0.1458**
Fwd-Look Dis. Accruals					**0.2621**

Panel C: Probit Analyses across Adjacent Earnings Change Groups

$$EM_{it} = \alpha + \beta_1 DisAccruals_{it} + \beta_2 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + \varepsilon_{it} \qquad (1)$$

Just-Beat Sample	Just-Miss Sample	n^{a}	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*ΔCash Flows*
$-.09 \leq \Delta E_{it} < -.08$	$-.10 \leq \Delta E_{it} < -.09$	813/677	−0.080^{c} (0.58)	0.390^{d} (0.44)	−0.029 (0.54)	−0.059^{d} (0.56)	0.062^{c} (0.43)
$-.08 \leq \Delta E_{it} < -.07$	$-.09 \leq \Delta E_{it} < -.08$	945/762	**0.462^{c}** **(0.10)**	2.340 (0.19)	**0.312** **(0.10)**	**0.651** **(0.05)**	−0.209^{c} (0.71)
$-.07 \leq \Delta E_{it} < -.06$	$-.08 \leq \Delta E_{it} < -.07$	1064/858	0.265^{c} (0.22)	2.045 (0.22)	−0.272^{d} (0.88)	−0.174^{c} (0.68)	0.304^{c} (0.16)
$-.06 \leq \Delta E_{it} < -.05$	$-.07 \leq \Delta E_{it} < -.06$	1272/1053	**0.817^{c}** **(0.01)**	−2.309^{c} (0.81)	**0.339** **(0.05)**	0.411 (0.12)	**0.443^{c}** **(0.09)**
$-.05 \leq \Delta E_{it} < -.04$	$-.06 \leq \Delta E_{it} < -.05$	1518/1245	−0.349^{c} (0.86)	**4.776** **(0.02)**	−0.060 (0.62)	−0.114^{c} (0.63)	0.060^{c} (0.43)
$-.04 \leq \Delta E_{it} < -.03$	$-.05 \leq \Delta E_{it} < -.04$	1827/1469	**0.925^{c}** **(0.00)**	0.424^{d} (0.42)	0.168 (0.18)	**0.518** **(0.06)**	**0.447^{c}** **(0.07)**
$-.03 \leq \Delta E_{it} < -.02$	$-.04 \leq \Delta E_{it} < -.03$	2284/1844	**0.550^{c}** **(0.03)**	**5.382** **(0.01)**	0.157 (0.17)	0.335 (0.13)	0.107^{c} (0.36)
$-.02 \leq \Delta E_{it} < -.01$	$-.03 \leq \Delta E_{it} < -.02$	2825/2311	**0.779^{c}** **(0.00)**	0.713^{d} (0.36)	0.179 (0.12)	**0.515** **(0.05)**	**0.762^{c}** **(0.00)**
$-.01 \leq \Delta E_{it} < .00$	$-.02 \leq \Delta E_{it} < -.01$	3539/2919	**0.963^{c}** **(0.00)**	**6.405** **(0.00)**	**0.409** **(0.00)**	**1.342** **(0.00)**	**0.861^{c}** **(0.00)**

(continued on next page)

TABLE 2 (Continued)

$.00 \leq \Delta E_{it} < .01$	$-.01 \leq \Delta E_{it} < .00$	4851/4028	**1.625** **(0.00)**	**4.442** **(0.01)**	0.088 (0.25)	**0.635** **(0.01)**	**1.883** **(0.00)**
$.01 < \Delta E_{it} < .02$	$.00 \leq \Delta E_{it} < .01$	5490/4636	**1.180** **(0.00)**	**2.259** **(0.09)**	**0.210** **(0.05)**	**1.176** **(0.00)**	**0.954**^{c} **(0.00)**
$.02 \leq \Delta E_{it} < .03$	$.01 \leq \Delta E_{it} < .02$	4461/3763	-0.192^{c} (0.76)	-1.544^{c} (0.81)	0.040 (0.39)	-0.232^{c} (0.80)	**0.469**^{c} **(0.03)**
$.03 \leq \Delta E_{it} < .04$	$.02 \leq \Delta E_{it} < .03$	3270/2748	**0.433**^{c} **(0.07)**	-0.151^{c} (0.53)	0.013 (0.47)	**0.568** **(0.03)**	**0.474**^{c} **(0.03)**
$.04 \leq \Delta E_{it} < .05$	$.03 \leq \Delta E_{it} < .04$	2427/2050	-0.215^{c} (0.76)	0.886^{d} (0.33)	−0.042 (0.59)	-0.124^{c} (0.66)	-0.012^{c} (0.53)
$.05 \leq \Delta E_{it} < .06$	$.04 \leq \Delta E_{it} < .05$	1868/1582	0.066^{c} (0.42)	-1.897^{c} (0.81)	0.014 (0.47)	**0.428** **(0.10)**	-0.267^{c} (0.84)
$.06 \leq \Delta E_{it} < .07$	$.05 \leq \Delta E_{it} < .06$	1560/1306	**0.857**^{c} **(0.01)**	2.120 (0.19)	−0.074 (0.63)	**0.760** **(0.02)**	**0.763**^{c} **(0.00)**
$.07 \leq \Delta E_{it} < .08$	$.06 \leq \Delta E_{it} < .07$	1269/1058	-0.249^{c} (0.74)	-2.409^{c} (0.82)	0.239 (0.15)	-0.386^{c} (0.84)	0.096^{c} (0.37)
$.08 \leq \Delta E_{it} < .09$	$.07 \leq \Delta E_{it} < .08$	1070/884	0.052^{c} (0.45)	-1.798^{c} (0.75)	-0.237^{d} (0.85)	0.173 (0.33)	-0.218^{c} (0.75)
$.09 \leq \Delta E_{it} < .10$	$.08 \leq \Delta E_{it} < .09$	908/753	0.330^{c} (0.22)	0.155^{d} (0.48)	0.293 (0.11)	0.344 (0.23)	-0.098^{c} (0.60)
$.10 \leq \Delta E_{it} < .11$	$.09 \leq \Delta E_{it} < .10$	781/648	-0.276^{c} (0.72)	-3.779^{c} (0.92)	−0.140 (0.70)	-0.099^{d} (0.58)	**0.500**^{c} **(0.08)**
Number of Positive Significant Coefficients for Pseudo Targets[b]			9 (0.00)	5 (0.00)	4 (0.00)	8 (0.00)	9 (0.00)

(continued on next page)

TABLE 2 (Continued)

Panel D: Pooled Probit Analyses Comparing the Actual Earnings Increase Benchmark to All other Pseudo Targets

$$EM_{it} = \alpha_1 + \beta_1 DisAccruals_{it} + \beta_2 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + Actual^*(\alpha_2 + \beta_3 DisAccruals_{it} + \beta_4 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it}) + \varepsilon_{it} \qquad (2)$$

	***DisAccruals* is defined as:**			
	Total Accruals	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*
DisAccruals	0.339	1.075	0.094	0.313
	(0.00)	(0.02)	(0.01)	(0.00)
ΔCashFlows	0.263	0.175	0.194	0.255
	(0.00)	(0.00)	(0.00)	(0.00)
*Actual*DisAccruals*	1.286	3.367	−0.005	0.322
	(0.00)	(0.03)	(0.97)	(0.13)
*Actual*ΔCashFlows*	1.980	1.358	1.364	1.628
	(0.00)	(0.00)	(0.00)	(0.00)
n	44,042	44,042	44,042	36,594

We construct probability levels for all statistical tests in Panels A, C, and D using one-tailed tests. Panels A and C highlight the "*actual*" just-beat and just-miss groups for profit benchmark with "boxed" text.

Bold text in Panel A indicates that the mean discretionary accruals value is significantly larger than the mean value of the preceding group ($p \leq .10$, one-tailed test). **Bold** text in Panel B highlights the significant Pearson correlation (($p \leq .10$, two-tailed test). **Bold** text in Panel C highlights significant coefficient for discretionary accruals measures ($p \leq .10$, one-tailed test).

[a] The number to the left of the slash "/" identifies the sample size for the analyses when *Total Accruals, Deferred Tax Expense,* and *Modified Jones Dis. Accruals* proxy for discretionary accruals. The number to the right of the slash "/" is the sample size for the analyses when *Fwd-Look Dis. Accruals* proxy for discretionary accruals.

[b] In the last two rows of Panel C we report (1) the number of significant positive coefficients (*k*) for each discretionary accrual measures ($p \leq .10$, one-tailed test) and (2) the probability of "drawing" *k* or more significant positive coefficients assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance.

[c] and [d] in Panel C indicate that the coefficient from the probit regression using the sample restricted to just-beat and just-miss firms around the *actual* benchmark is significantly larger than the estimated coefficient for this specific *pseudo* target at $p \leq .05$ (one-tailed test) and $p \leq .10$ (one-tailed test), respectively.

Variable Definitions:

ΔE_{it} = the change in net income (annual Compustat data item #172) from year $t-1$ to year t scaled by the market value of equity at the end of year $t-2$ (annual Compustat data item #25 × #199);

EM_{it} = 1 if $X \leq \Delta E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \leq \Delta E_{it} < X$. X is the earnings change target. X equals 0 for the *actual* earnings increase benchmark; and

ΔCashFlows = the change in *Cash Flows* from year $t-1$ to year t, scaled by total assets as of the end of year $t-1$;

Table 1 provides the definitions for all other variables.

discretionary accrual proxies and reporting more positive earnings changes may not be unique to comparisons centered on the earnings increase benchmark.

Comparisons across Pseudo Targets

Table 2, Panel C reports the results of the *pseudo* target and *actual* benchmark analyses.[12] Each row represents the results for estimating Equation (1) for a different just-beat and just-miss sample. We highlight positive and significant coefficients ($p \le .10$, one-tailed test) with bold text and identify the *actual* just-beat and just-miss groups with boxed text. In the last two rows we tally the number of significant positive coefficients (*k*) for the *pseudo* target regressions for each discretionary accrual measure ($p \le .10$, one-tailed test) and present the probability of "drawing" *k* or more significant positive coefficients assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance. Panel C reports significant positive associations between (1) beating the *actual* earnings increase benchmark and (2) *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals* ($p \le .01$, one-tailed test). The coefficient for *Modified Jones Dis. Accruals* is positive but not statistically significant ($p = .25$, one-tailed test).

Results for the *pseudo earnings change* target analysis are similar to those for the *pseudo* profit target analysis, albeit less pronounced. Of the 19 comparisons across adjacent earnings change bins, segregated based on *pseudo earnings change* targets, we find a significant positive association ($p \le .10$, one-tailed test) between beating the *pseudo* target and (1) *Total Accruals* in nine comparisons (47 percent), (2) *Deferred Tax Expense* in five comparisons (26 percent), (3) *Modified Jones Dis. Accruals* in four comparisons (21 percent), and (4) *Fwd-Look Dis. Accruals* in eight comparisons (42 percent). These frequencies of significant positive coefficients are significantly higher than expected by chance at $p = .00$, $p = .04$, and $p = .00$ (one-tailed test) for *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals*, respectively, and at $p = .12$ (one-tailed test) for *Modified Jones Dis. Accruals*. Thus, similar to the profit benchmark analysis, it is difficult to distinguish whether the associations between *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals* and beating the earnings increase benchmark are attributable to earnings management or the associations between these discretionary accruals proxies and earnings changes. Finally, Table 2, Panel C indicates that *ΔCashFlows* is positively related to beating the earnings increase benchmark and nine of 19 *pseudo* earnings change targets ($p \le .10$, one-tailed test).

Tests for Shifts in the Association between Discretionary Accrual Measures and Firm Performance around the Earnings Increase Benchmark

We estimate Equation (2) to test whether the associations between discretionary accrual proxies and firm performance increase around the *actual* earnings increase benchmark (i.e., where earnings management incentives may be stronger). In Table 2, Panel C, we identify with superscripts "c" and "d" those regression coefficients from the *pseudo* target analyses that are significantly *smaller* ($p \le .05$ and $p \le .10$, one-tailed test, respectively) than the estimated coefficient from the *actual* earnings increase benchmark (i.e., significantly smaller than the coefficients highlighted in boxed text). In Table 2, Panel D, we report the results when we compare the associations between discretionary accrual proxies and beating the *actual* earnings increase benchmark to the *average* association between the discretionary accrual proxies and beating the *pseudo* targets.

---

[12] Our probit regression results comparing just-beat and just-miss firms around the *actual* earnings increase benchmark are similar to the probit results reported in Phillips et al. (2003).

Consistent with the relationship between discretionary accrual proxies and firm performance becoming more pronounced around the earnings increase benchmark, Table 2, Panel D reports that the regression coefficients for *Total Accruals* and *Deferred Tax Expense* estimated around the *actual* earnings increase benchmark are significantly larger ($p = .00$ and .03, respectively, one-tailed test) than the *average* association between the respective discretionary accrual measures and beating the *pseudo* earnings change targets. The difference between the coefficient for *Fwd-Look Dis. Accruals* estimated around the earnings increase benchmark and the average association between *Fwd-Look Dis. Accruals* and beating the *pseudo* earnings change targets is positive as expected but not statistically significant ($p = .13$, one-tailed test).

Table 2, Panel C reports that the coefficients for *Total Accruals*, *Deferred Tax Expense*, and *Fwd-Look Dis. Accruals* from the regressions estimated around the *actual* earnings increase benchmark are significantly larger than many of the estimated coefficients from the *pseudo* profit target comparisons. In particular, the estimated coefficients around the earnings increase benchmark are significantly larger than 18 of the 19 coefficients for *Total Accruals* for the *pseudo* targets, 12 of the 19 coefficients for *Deferred Tax Expense* for the *pseudo* targets, and eight of the 19 coefficients for *Fwd-Look Dis. Accruals* for the *pseudo* targets ($p < .10$, one-tailed test). These frequencies are significantly higher than expected ($p = .00$, one-tailed test) assuming a binomial distribution and 10 percent likelihood that the *actual* benchmark coefficient is significantly larger than a *pseudo* target coefficient by chance.

The associations between Δ*CashFlows* and reporting more positive earnings changes generally also increase around the *actual* earnings increase benchmark. Specifically, Table 2, Panel D reports that the association between Δ*CashFlows* and beating the *actual* earnings increase benchmark is significantly larger ($p \leq =.00$, one-tailed test) than the *average* association between Δ*CashFlows* and beating the *pseudo* earnings change targets in all four specifications of Equation (2). Likewise, Table 2, Panel C indicates that the estimated coefficient for Δ*CashFlows* around the *actual* earnings increase benchmark is significantly larger than each of the coefficients for Δ*CashFlows* for the *pseudo* earnings change targets. As discussed above, these results are consistent with firms managing both cash flows and discretionary accruals around the earnings increase benchmark, or managers and employees working harder around the earnings increase benchmark. Accordingly, we are unable to conclude with confidence that earnings management explains the associations between discretionary accruals and beating the earnings increase benchmark.

Analysts' Forecast Benchmark Analyses

Table 3, Panels A and B report the univariate analyses for the analysts' forecast benchmark. Panel A indicates that *Total Accruals*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are significantly larger (more positive or less negative) in the *actual* just-beat bin relative to the *actual* just-miss bin.[13] These same relations hold for only one of the 19 other comparisons for *Modified Jones Dis. Accruals* and none of the 19 other comparisons for *Total Accruals* and *Fwd-Look Dis. Accruals* ($p \leq .10$, one-tailed test). Unlike the results for the profit and earnings increase benchmark analyses, there appears to be little positive

---

[13] Using the last mean consensus analyst forecast from the Adjusted I/B/E/S Summary file as the analysts' forecast benchmark, prior research (Dechow et al. 2000; Phillips et al. 2003) reports no significant differences in univariate comparisons of *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* for just-beat and just-miss firms around the *actual* analysts' forecast benchmark. We find similar results to prior research when we use the last mean consensus analyst forecast from the Adjusted I/B/E/S Summary file as the analysts' forecast benchmark.

TABLE 3
Analysts' Forecast Benchmark Analyses

Panel A: Descriptive Statistics

UE_{it}	n^a	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*ΔCash Flows*
$UE_{it} = -.10$	151/132	−0.0660	−0.0031	−0.0065	0.0036	0.0018
$UE_{it} = -.09$	205/166	−0.0585	**−0.0007**	−0.0115	0.0079	0.0062
$UE_{it} = -.08$	243/199	−0.0558	0.0003	0.0022	0.0100	−0.0170
$UE_{it} = -.07$	234/190	−0.0477	−0.0017	0.0030	0.0167	0.0006
$UE_{it} = -.06$	268/227	−0.0483	−0.0011	0.0043	0.0009	0.0181
$UE_{it} = -.05$	366/297	−0.0467	−0.0005	0.0026	0.0054	−0.0073
$UE_{it} = -.04$	380/308	−0.0557	−0.0003	−0.0060	0.0008	0.0019
$UE_{it} = -.03$	523/428	−0.0549	**0.0010**	0.0041	0.0100	0.0100
$UE_{it} = -.02$	722/585	−0.0513	−0.0001	−0.0008	0.0089	0.0116
$UE_{it} = -.01$	1109/866	−0.0500	**0.0011**	−0.0033	0.0055	0.0178
$UE_{it} = .00$	2200/1757	**−0.0423**	0.0016	**0.0099**	**0.0142**	0.0157
$UE_{it} = .01$	2045/1644	−0.0410	0.0012	0.0126	0.0138	**0.0207**
$UE_{it} = .02$	1639/1308	−0.0476	0.0012	0.0018	0.0080	**0.0269**
$UE_{it} = .03$	1226/985	−0.0477	0.0016	0.0050	0.0093	0.0205
$UE_{it} = .04$	880/728	−0.0487	0.0005	0.0007	0.0098	0.0210
$UE_{it} = .05$	718/588	−0.0504	0.0002	−0.0062	0.0034	0.0249
$UE_{it} = .06$	503/425	−0.0544	−0.0007	−0.0063	−0.0016	0.0253
$UE_{it} = .07$	436/350	−0.0569	0.0001	0.0032	−0.0041	0.0136
$UE_{it} = .08$	351/298	−0.0603	−0.0012	−0.0233	0.0016	0.0181
$UE_{it} = .09$	272/210	−0.0550	−0.0008	**−0.0075**	0.0054	0.0153
$UE_{it} = .10$	222/187	−0.0532	−0.0006	0.0070	0.0017	0.0268
Pearson Correlation between Mean Discretionary Accrual Measures and Bin Number		0.09	0.26	−0.09	−0.36	0.73
(p-value, two-tailed)		(.69)	(.26)	(.70)	(.11)	(.00)

(*continued on next page*)

TABLE 3 (Continued)

Panel B: Pearson Correlations

	Total Accruals	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*ΔCash Flows*
UE	−0.0111	0.0004	−0.0147	**−0.0153**	**0.0627**
Total Accruals		0.0087	**0.4399**	**0.6469**	**−0.3620**
Deferred Tax Expense			**0.0449**	**0.0545**	**0.0265**
Modified Jones Dis. Accruals				**0.4370**	**−0.1962**
Fwd-Look Dis. Accruals					**−0.3425**

Panel C: Probit Analyses across Adjacent Unexpected Earnings Groups

$$EM_{it} = \alpha + \beta_1 DisAccruals_{it} + \beta_2 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + \varepsilon_{it} \quad (1)$$

Just-Beat Sample	**Just-Miss Sample**	**n**[a]	*Total Accruals*	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*	*ΔCash Flows*
$UE_{it} = -.09$	$UE_{it} = -.10$	354/293	−0.000 (0.50)	4.755 (0.17)	−0.200 (0.67)	−0.136 (0.58)	0.123 (0.43)
$UE_{it} = -.08$	$UE_{it} = -.09$	445/363	0.037 (0.48)	0.494 (0.45)	0.313 (0.23)	−0.396^{d} (0.72)	−0.615 (0.92)
$UE_{it} = -.07$	$UE_{it} = -.08$	473/386	0.435 (0.24)	−5.591^{d} (0.91)	−0.121 (0.60)	0.371 (0.30)	0.516 (0.13)
$UE_{it} = -.06$	$UE_{it} = -.07$	498/412	0.274 (0.33)	1.287 (0.38)	−0.135 (0.62)	−0.725^{d} (0.83)	0.584 (0.18)
$UE_{it} = -.05$	$UE_{it} = -.06$	625/514	−0.194 (0.64)	**5.830** **(0.05)**	−0.223^{d} (0.74)	0.299 (0.32)	−0.868^{d} (0.93)
$UE_{it} = -.04$	$UE_{it} = -.05$	742/601	−0.561^{c} (0.88)	0.185 (0.48)	−0.141 (0.68)	−0.026 (0.52)	0.275 (0.31)
$UE_{it} = -.03$	$UE_{it} = -.04$	893/728	0.466 (0.16)	2.307 (0.24)	**0.586** **(0.02)**	0.662 (0.11)	0.770 (0.05)
$UE_{it} = -.02$	$UE_{it} = -.03$	1235/1004	−0.016 (0.52)	−4.072^{c} (0.94)	−0.203^{d} (0.78)	−0.420^{c} (0.83)	−0.098 (0.60)
$UE_{it} = -.01$	$UE_{it} = -.02$	1817/1447	0.223 (0.26)	**3.063** **(0.06)**	−0.181^{c} (0.80)	−0.066^{d} (0.57)	0.381 (0.13)

(continued on next page)

TABLE 3 (Continued)

UE_{it} = .00	UE_{it} = −.01	3285/2592	**0.434** **(0.06)**	0.876 (0.29)	**0.270** **(0.04)**	**0.547** **(0.04)**	0.006 (0.49)
UE_{it} = .01	UE_{it} = .00	4228/3378	**0.347** **(0.07)**	−0.879 (0.75)	0.047 (0.35)	0.263 (0.16)	**0.425** **(0.04)**
UE_{it} = .02	UE_{it} = .01	3664/2936	−0.373^{c} (0.93)	−0.008 (0.50)	−0.189^{c} (0.92)	−0.342^{c} (0.89)	0.103 (0.34)
UE_{it} = .03	UE_{it} = .02	2856/2278	−0.029 (0.54)	0.553 (0.36)	0.103 (0.26)	0.192 (0.27)	−0.153 (0.71)
UE_{it} = .04	UE_{it} = .03	2100/1704	−0.255^{d} (0.78)	−3.428^{c} (0.97)	−0.117^{d} (0.74)	0.174 (0.31)	0.071 (0.41)
UE_{it} = .05	UE_{it} = .04	1583/1303	0.096 (0.40)	−0.359 (0.57)	−0.112^{d} (0.70)	−0.405^{c} (0.85)	−0.169 (0.68)
UE_{it} = .06	UE_{it} = .05	1213/1009	−0.269^{d} (0.74)	−1.751 (0.77)	0.130 (0.29)	−0.235^{d} (0.70)	0.058 (0.44)
UE_{it} = .07	UE_{it} = .06	933/770	−0.425^{d} (0.79)	0.773 (0.39)	0.193 (0.24)	−0.163 (0.62)	−0.519 (0.87)
UE_{it} = .08	UE_{it} = .07	782/641	−0.366^{d} (0.75)	−4.735^{c} (0.94)	−0.705^{c} (0.99)	−0.015 (0.51)	0.342 (0.26)
UE_{it} = .09	UE_{it} = .08	618/502	0.407 (0.25)	0.422 (0.45)	0.396 (0.14)	0.284 (0.33)	−0.085 (0.56)
UE_{it} = .10	UE_{it} = .09	488/391	−0.045 (0.52)	−1.534 (0.64)	0.362 (0.16)	−0.349 (0.69)	0.438 (0.21)
Number of Positive Significant Coefficients for Pseudo Targets[b]			1 (0.87)	2 (0.58)	1 (0.87)	0 (1.00)	1 (0.87)

(continued on next page)

TABLE 3 (Continued)

Panel D: Pooled Probit Analyses Comparing the Actual Analysts' Forecast Benchmark to All other Pseudo Targets

$$EM_{it} = \alpha_1 + \beta_1 DisAccruals_{it} + \beta_2 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it} + Actual^*(\alpha_2 + \beta_3 DisAccruals_{it} + \beta_4 \Delta CashFlows_{it} + \beta_j \Sigma_j Ind_{it}) + \varepsilon_{it} \qquad (2)$$

	***DisAccruals* is defined as:**			
	Total Accruals	*Deferred Tax Expense*	*Modified Jones Dis. Accruals*	*Fwd-Look Dis. Accruals*
DisAccruals	−0.008	−0.089	−0.024	−0.021
	(0.54)	(0.57)	(0.68)	(0.58)
ΔCashFlows	0.088	0.091	0.085	0.089
	(0.11)	(0.09)	(0.11)	(0.15)
*Actual*DisAccruals*	0.442	0.965	0.294	0.567
	(0.07)	(0.28)	(0.03)	(0.04)
*Actual*ΔCashFlows*	−0.070	−0.190	−0.107	−0.083
	(0.62)	(0.81)	(0.68)	(0.61)
n	28,832	28,832	28,832	23,252

We construct probability levels for all statistical tests in Panels A, C, and D using one-tailed tests. Panels A and C highlight the *actual* just-beat and just-miss groups for profit benchmark with "boxed text."

Bold text in Panel A indicates that the mean discretionary accruals value is significantly larger than the mean value of the preceding group ($p \leq .10$, one-tailed test). **Bold** text in Panel B highlights the significant Pearson correlation (($p \leq .10$, two-tailed test). **Bold** text in Panel C highlights significant coefficient for discretionary accruals measures ($p \leq .10$, one-tailed test).

[a] The number to the left of the slash "/" identifies the sample size for the analyses when *Total Accruals*, *Deferred Tax Expense*, and *Modified Jones Dis. Accruals* proxy for discretionary accruals. The number to the right of the slash "/" is the sample size for the analyses when *Fwd-Look Dis. Accruals* proxy for discretionary accruals.

[b] In the last two rows of Panel C we report (1) the number of significant positive coefficients (*k*) for each discretionary accrual measures ($p \leq .10$, one-tailed test) and (2) the probability of "drawing" *k* or more significant positive coefficients assuming a binomial distribution and 10 percent likelihood of drawing a significant positive coefficient by chance.

[c] or [d] in Panel C indicate that the coefficient from the probit regression using the sample restricted to just-beat and just-miss firms around the *actual* benchmark is significantly larger than the estimated coefficient for this specific *pseudo* target at $p \leq .05$ (one-tailed test) or $p \leq .10$ (one-tailed test), respectively.

Variable Definitions:

UE_{it} = the unexpected earnings (actual earnings per share minus the single-most recent analyst forecast provided prior to the earnings announcement both available from the Unadjusted I/B/E/S Detail History file); and

EM_{it} = 1 if UE_{it} equals *X* cents per share and 0 if UE_{it} equals *X* less one cent per share. We define UE_{it} as firm *i*'s year *t* actual earnings per share minus the single-most recent analyst forecast provided prior to the earnings announcement both reported from the Unadjusted I/B/E/S Detail History file. We round UE_{it} to the nearest penny. We define *X* as the unexpected earnings target, and *X* equals zero for the *actual* analysts' forecast benchmark.

Table 1 provides the definitions for all other variables.

correlation between the mean values of the discretionary accrual measures and the unexpected earnings bins. In particular, the Pearson correlation coefficients for *Total Accruals*, *Deferred Tax Expense*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* are .09, .26, −.09, and −.36, respectively. None of the correlations are statistically significant. Likewise, Panel B reports that none of four discretionary accrual proxies have a significant positive correlation with *UE*.

Comparisons across Pseudo Targets

Table 3, Panel C reports a significant positive association between (1) beating the *actual* analysts' forecast benchmark and (2) *Total Accruals*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* ($p = .06$, .04, and .04, respectively, one-tailed test).[14] We find a positive but insignificant association for *Deferred Tax Expense* ($p = .29$, one-tailed test). In addition, we find few positive associations between the discretionary accrual measures and beating the *pseudo* analysts-based unexpected earnings targets. Specifically, the frequency of significant positive coefficients for each discretionary accrual measures is not significantly different from what would be expected by chance. Finally, Panel C indicates that Δ*CashFlows* is not significantly associated with beating the analysts' forecast benchmark and is only significantly positively related to beating one of the 19 *pseudo* unexpected earnings targets ($p \leq .10$, one-tailed test).[15]

Tests for Shifts in the Association between Discretionary Accrual Measures and Firm Performance around the Analysts' Forecast Benchmark

Because we find little evidence of a positive association between discretionary accrual measures and beating *pseudo* targets derived from analysts-based unexpected earnings, the comparisons of the coefficients for *Total Accruals*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* across the *actual* analysts' forecast benchmark and *pseudo* targets are less informative. Nonetheless, Table 3, Panel D reports that the coefficients for *Total Accruals*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* estimated around the *actual* analysts' forecast benchmark are significantly larger ($p = .07$, .03, and .04, respectively, one-tailed test) than the *average* association between the respective discretionary accrual measures and beating the *pseudo* analysts' forecast targets. Likewise, Panel C reports that the coefficients for *Total Accruals*, *Modified Jones Dis. Accruals*, and *Fwd-Look Dis. Accruals* estimated around the *actual* analysts' forecast benchmark are significantly larger than many of the estimated coefficients from the *pseudo* profit target comparisons. Finally, unlike the profit and earnings increase benchmark analyses, we find no associations between Δ*CashFlows* and beating the analysts' forecast benchmark. Accordingly, we are able to draw stronger conclusions regarding earnings management around the analysts' forecast benchmark.

---

[14] Phillips et al. (2003) uses the last mean consensus analyst forecast from the Adjusted I/B/E/S Summary file as the analysts' forecast benchmark. When we use the last mean consensus analyst forecast from the Adjusted I/B/E/S Summary file, probit results for the *actual* analysts' forecast benchmark are similar to those reported in Phillips et al. (2003).

[15] In sensitivity analysis, we compared the price-to-earnings (PE) multiples and book-to-market (BM) ratios of firms that just beat the analysts' forecast benchmark to other sample firms to determine if the just-beat firms are more likely to represent growth firms. We find no significant difference in PE multiples or book-to-market ratios for firms that just beat the analysts' forecast benchmark (i.e., $UE = .00$) compared to those firms that just miss the benchmark ($UE = -.01$).

Alternative Analysts' Forecast Specifications

We use the single-most recent forecast from the Unadjusted I/B/E/S Detail History file as the forecast proxy in our main analyses. O'Brien (1988) and Brown (1991) suggest that the single-most recent analyst forecast is more accurate in predicting actual earnings than the consensus mean forecast. Likewise, Brown and Kim (1991) find that the single-most recent analyst forecast more accurately reflects the market's earnings expectation than the consensus mean forecast (i.e., the average of all outstanding forecasts). I/B/E/S recently released the Unadjusted Detail History file that includes analysts' forecasts that have not been adjusted for subsequent stock splits. Previously, I/B/E/S only provided earnings forecasts adjusted for subsequent stock splits, which prior research indicates may result in classification errors due to rounding (Baber and Kang 2002; Payne and Thomas 2003).

In sensitivity analysis, we re-estimate our probit regressions to determine whether results are robust to other forecast measures or are sensitive to using data from the Unadjusted or Adjusted I/B/E/S files. We first re-estimate our analyses using data from the Adjusted I/B/E/S Detail History file (1) including all sample observations, (2) eliminating firms with a stock split factor exceeding 3, and (3) eliminating firms with a stock split factor exceeding 1. Results are consistent with those presented in Table 3, Panels C and D for each alternative sample.

We also re-estimate our analyses using the last mean consensus analyst forecast from the Unadjusted and Adjusted I/B/E/S Summary files. Results weaken when we use the last mean consensus analyst forecast from the Unadjusted file and disappear when we use the last mean consensus analyst forecast from the Adjusted file. Specifically, using data from the Unadjusted I/B/E/S Summary file, we find that only two of the abnormal accrual measures (i.e., *Modified Jones Dis. Accruals* and *Fwd-Look Dis. Accruals*) are statistically significant ($p = .09$, one-tailed test). When we use the last mean consensus forecast reported by the Adjusted I/B/E/S Summary file (1) including all sample observations, (2) eliminating firms with a stock split factor exceeding 3, and (3) eliminating firms with a stock split factor exceeding 1, we find that none of the discretionary accrual measures are positively associated with beating the analysts' forecast benchmark. These results suggest that the analysts' forecast results weaken when we use a more "stale" forecast proxy. Assuming that firms intend to meet or beat market expectations, one would expect that results improve when we use a forecast proxy that better represents these expectations.

V. SUPPLEMENTAL ANALYSES

ROA Performance-Matching

Kothari et al. (2005) provide evidence that "performance matching is critical to designing well-specified tests of earnings management." In particular, they find that ROA performance-matched accrual measures mitigate Type I research errors when the partitioning variable is correlated with performance. One advantage of our analysis relative to performance-matching is that it does not require specifying the exact nature of the correlation between discretionary accruals and performance—i.e., it does not require specifying a metric to match upon (e.g., ROA). Accordingly, relative to performance-matching this analysis may result in more powerful tests of earnings management around earnings benchmarks as it does not introduce "noise" in the tests via the matching process.

To compare our results to those from performance-matching, we re-estimate our probit regressions using ROA performance-matched discretionary accrual measures. For the earnings change and analysts-based unexpected earnings analyses, we match firms based on their current ROA and two-digit industry as suggested by Kothari et al. (2005). Because comparisons based on firm profits directly relate to current ROA, we match firms on lagged ROA and two-digit SIC industry for the profit target analyses.[16] Kothari et al. (2005) find that matching on current or lagged ROA results in better-specified tests than discretionary accrual measures that are not performance-matched, although matching on current ROA generally performs better than lagged ROA matching. We then re-estimate our probit models using discretionary accrual measures adjusted by the matched firm's discretionary accruals.

Table 4 presents the results including *Fwd-Look Dis. Accruals* as the discretionary accrual proxy in the profit, earnings increase, and analysts' forecast benchmark analyses. Results are generally much weaker than those reported in Panel C of Tables 1, 2, and 3. In particular, we find that *Fwd-Look Dis. Accruals* is no longer significantly associated with beating the *actual* profit or analysts' forecast benchmark. Conclusions are similar when we use *Total Accruals*, *Deferred Tax Expense*, or *Modified Jones Dis. Accruals* to proxy for discretionary accruals—i.e., we find little evidence of significant positive associations between performance-matched discretionary accrual proxies and beating the profit or analysts' forecast benchmark.

There are two plausible explanations for the considerably weaker results for performance-matched discretionary accruals: (1) performance-matching reduces the power of statistical tests of earnings management around earnings benchmarks, and (2) the previous associations we find between discretionary accrual measures and beating the *profit* and *analysts' forecast* benchmarks are attributable to an underlying association between discretionary accrual proxies and firm performance that is unrelated to earnings management. Our previous analyses indicate that discretionary accrual proxies are not systematically associated with reporting more positive analysts-based unexpected earnings. On balance, therefore, we interpret explanation (1) to be more descriptive for the performance-matching results for the analysts' forecast benchmark. Given our prior evidence of a systematic positive association between discretionary accrual proxies and reporting higher earnings, we are unable to distinguish explanations (1) and (2) for the performance-matching results for the profit benchmark.

Alternative Explanations for the Profit and Earnings Change Kinks

Recent studies provide alternative explanations (i.e., besides earnings management) for the kinks in the earnings and earnings changes distributions around the profit and earnings increase benchmarks. Dechow et al. (2003) and Durtschi and Easton (2005) demonstrate that the kinks in the earnings and earnings change distributions disappear when earnings and earnings changes are deflated by shares outstanding instead of market value. In contrast, Beaver et al. (2003) contend that much of the kink in the earnings and earnings changes distributions are attributable to factors not related to earnings management, such as the asymmetric treatment of income taxes and special items for loss firms versus profitable firms.

We conduct additional analyses to control for the effects that the asymmetric treatment of income taxes and special items for loss and profits may have on our inferences. Specifically, we re-estimate the profit, earnings increase, and analysts' forecast benchmark analyses after including two additional control variables, special items deflated by lagged assets and current tax expense deflated by lagged assets. Results are similar to those presented in the tables. In sum, we find no evidence that differences in the asymmetric treatment of

---

[16] Consistent with Kothari et al. (2005), we match sample firms to the firm in their two-digit industry with the closest current or lagged ROA. In addition, we require that the matched firm resides in a bin other than the just-beat or just-miss bin for the specific comparison.

TABLE 4
Probit Analyses of ROA Performance-Matched Forward Look Discretionary Accruals across Adjacent Profit, Earnings Change, Unexpected Earnings Groups

$$EM_{it} = \alpha + \beta_1 ROA\text{-}Matched\ Fwd\text{-}Look\ Dis.\ Accruals_{it} + \beta_2 \Delta Cash\ Flows_{it} + \beta_j\ \Sigma_j\ Ind_{it} + \varepsilon_{it} \quad (1)$$

		Profit		Earnings Change		Unexpected Earnings	
Just-Beat Sample[a]	Just-Miss Sample[b]	n	*ROA-Matched Fwd-Look Dis. Accruals*	n	*ROA-Matched Fwd-Look Dis. Accruals*	n	*ROA-Matched Fwd-Look Dis. Accruals*
$-.09 \leq E_{it} < -.08$	$-.10 \leq E_{it} < -.09$	654	−0.131 (0.24)	677	0.234 (0.19)	293	−0.594 (0.23)
$-.08 \leq E_{it} < -.07$	$-.09 \leq E_{it} < -.08$	679	0.001 (0.50)	762	0.044 (0.43)	363	0.182 (0.69)
$-.07 \leq E_{it} < -.06$	$-.08 \leq E_{it} < -.07$	707	0.179 (0.16)	858	−0.074 (0.38)	386	0.305 (0.52)
$-.06 \leq E_{it} < -.05$	$-.07 \leq E_{it} < -.06$	803	−0.160 (0.17)	1053	−0.166 (0.22)	412	**1.041** **(0.04)**
$-.05 \leq E_{it} < -.04$	$-.06 \leq E_{it} < -.05$	878	0.018 (0.46)	1245	**0.305** **(0.07)**	514	−0.802 (0.06)
$-.04 \leq E_{it} < -.03$	$-.05 \leq E_{it} < -.04$	941	0.011 (0.48)	1469	0.197 (0.17)	601	−0.768 (0.04)
$-.03 \leq E_{it} < -.02$	$-.04 \leq E_{it} < -.03$	1013	0.158 (0.18)	1844	0.128 (0.23)	728	0.146 (0.674)
$-.02 \leq E_{it} < -.01$	$-.03 \leq E_{it} < -.02$	1101	−0.094 (0.30)	2311	0.085 (0.30)	1004	0.133 (0.65)
$-.01 \leq E_{it} < .00$	$-.02 \leq E_{it} < -.01$	1107	0.104 (0.27)	2919	0.064 (0.32)	1447	−0.009 (0.97)
$.00 \leq E_{it} < .01$	$-.01 \leq E_{it} < .00$	1494	0.007 (0.48)	4029	**0.184** **(0.06)**	2592	0.032 (0.86)
$.01 < E_{it} < .02$	$.00 \leq E_{it} < .01$	1996	0.162 (0.12)	4636	**0.228** **(0.02)**	3378	0.117 (0.44)
$.02 \leq E_{it} < .03$	$.01 \leq E_{it} < .02$	2266	0.065 (0.31)	3763	0.030 (0.41)	2936	−0.171 (0.30)

(continued on next page)

TABLE 4 (Continued)

$.03 \le E_{it} < .04$	$.02 \le E_{it} < .03$	2667	0.051 (0.35)	2748	0.088 (0.29)	2278	−0.047 (0.80)
$.04 \le E_{it} < .05$	$.03 \le E_{it} < .04$	3089	**0.266** **(0.02)**	2050	−0.172 (0.16)	1704	0.235 (0.28)
$.05 \le E_{it} < .06$	$.04 \le E_{it} < .05$	3512	−0.025 (0.42)	1582	−0.031 (0.44)	1303	0.093 (0.70)
$.06 \le E_{it} < .07$	$.05 \le E_{it} < .06$	3655	−0.167 (0.07)	1306	0.238 (0.13)	1009	0.113 (0.69)
$.07 \le E_{it} < .08$	$.06 \le E_{it} < .07$	3328	**0.203** **(0.06)**	1058	0.251 (0.14)	770	−0.172 (0.60)
$.08 \le E_{it} < .09$	$.07 \le E_{it} < .08$	2758	0.069 (0.32)	884	−0.352 (0.08)	641	−0.043 (0.91)
$.09 \le E_{it} < .10$	$.08 \le E_{it} < .09$	2183	−0.112 (0.26)	753	0.037 (0.45)	502	0.026 (0.95)
$.10 \le E_{it} < .11$	$.09 \le E_{it} < .10$	1758	−0.029 (0.44)	648	−0.038 (0.45)	391	0.328 (0.52)

---

[a] For analysts-based unexpected earnings bins, just-beat firms include firms for which unexpected earnings equal the lower bound value.

[b] For analysts-based unexpected earnings bins, just-miss firms include firms for which unexpected earnings equal the lower bound value.

We construct probability levels for all coefficients using a one-tailed test. We highlight the *actual* just-beat and just-miss groups for a specific earnings benchmark with "boxed text" and significant positive coefficients ($p \le .10$, one-tailed test) with **bold** text.

In the profit benchmark analyses, EM_{it} equals 1 if $X \le E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \le E_{it} < X$. We define E_{it} as firm i's net income in year t (annual Compustat data item #172) divided by the market value of equity at the end of year $t-1$ (annual Compustat data item #25 × #199), and X as the profit target. X equals 0 for the *actual* profit benchmark.

In the earnings change benchmark analyses, EM_{it} equals 1 if $X \le \Delta E_{it} < (X + 0.01)$ and 0 if $(X - 0.01) \le \Delta E_{it} < X$. We define ΔE_{it} as the change in firm i's net income from year $t-1$ to t divided by the market value of equity at the end of year $t-2$, and X as the earnings change target. X equals 0 for the *actual* earnings increase benchmark.

In the analysts-based unexpected earnings target analyses, EM_{it} equals 1 if UE_{it} equals X and 0 if UE_{it} equals X less one cent per share. We define UE_{it} as firm i's year t actual earnings per share minus the single-most recent analyst forecast provided prior to the earnings announcement both reported from the Unadjusted I/B/E/S Detail History file. We round UE_{it} to the nearest penny. We define X as the unexpected earnings target, and X equals 0 for the *actual* analysts' forecast benchmark.

To calculate *ROA-Matched Fwd-Look Dis. Accruals*, we match firms on ROA and two-digit SIC industry (Kothari et al. 2005).

Table 1 provides the definitions for all other variables.

special items and income taxes explain the previous associations between discretionary accruals and beating the profit, earnings increase, and analysts' forecast benchmarks.

AAER Analyses

We collected a sample of firms subject to SEC enforcement action for overstating earnings to determine if these firms are more likely to reside in the just-beat bins around the *actual* earnings benchmarks. Following the method in Dechow et al. (1995, 1996), we conducted a keyword search for "Section 13(a)" in the Accounting and Auditing Enforcement Releases (AAERs) Database through Lexis-Nexis. We searched for AAERs issued between 1994 and November 2004 with an identified manipulation between fiscal year 1994 and 2002 to match our sample period. We eliminated releases involving violation of auditing standards, multiple releases involving the same firms, and firms not listed on Compustat. The final sample consisted of 392 firm-years subject to enforcement actions by the SEC for overstating earnings between 1994 and 2002.

We then plotted the percentage of AAER firm-years (i.e., the violation years) in each "bin" of the distribution of earnings, earnings changes, and analysts-based unexpected earnings against the relative percentages for this study's sample of firms (i.e., from Panel A of Tables 1, 2 and 3). The percentage of AAER firms in the $UE = 0$ bin for the analysts' forecast benchmark (39.7 percent) (not tabulated) is significantly higher (p = .00, one-tailed binomial test) than the underlying population of firms that reside in this bin (15.0 percent). Less pronounced differences exist for the profit benchmark (i.e., 6.9 percent of AAER firms reside in the $.00 \leq E_{it} < .01$. bin versus 5.2 percent for all firms, p = .14, one-tailed binomial test). For the earnings increase benchmark, the percentages of AAER firms in the $.00 \leq \Delta E_{it} < .01$ bin (8.8 percent) is *smaller* than the underlying population of firms that reside in this bin (13.0 percent). Similar to previous analyses, these results allow stronger conclusions regarding earnings management around the analysts' forecast benchmark. This evidence is also consistent with recent studies that conclude that the analysts' forecast benchmark is the most important benchmark during the sample period (Dechow et al. 2003; Brown and Caylor 2005).

VI. CONCLUSION

This study investigates the extent to which the positive associations between alternative discretionary accrual measures and beating benchmarks are attributable to earnings management. We find that the positive associations between discretionary accrual measures and beating the profit benchmark hold for comparisons of groups segregated based on *pseudo* profit targets. We find similar results for *pseudo* targets across the earnings change distribution. In contrast, we find that the positive associations between discretionary accruals and beating the analysts' forecast benchmark generally do *not* extend to comparisons across *pseudo* targets based upon analysts-based unexpected earnings. These results suggest that demonstrating positive associations between a specific discretionary accrual measure and beating the profit and earnings increase benchmark is not sufficient to conclude that the discretionary accrual actually detects earnings management. Likewise, this evidence indicates that the underlying association between firm performance and discretionary accrual measures is less problematic for tests of earnings management using analysts-based unexpected earnings.

We provide an additional analysis that specifically accounts for the systematic association between discretionary accrual measures and earnings. Implicit in studies analyzing accruals around earnings benchmarks is the expectation that the fundamental relationship between firm performance and discretionary accruals shifts around the earnings benchmarks. Accordingly, we test whether the associations between discretionary accrual proxies and firm performance are more pronounced around the profit benchmark. We conduct similar analyses for the *actual* earnings increase benchmark. Consistent with earnings management, results suggest that the association between discretionary accruals and earnings intensifies around the profit benchmark. We find similar effects around the *actual* earnings increase benchmark. However, analogous patterns exist for cash flows around the profit and earnings increase benchmarks. Thus, we are unable to conclude with confidence that earnings management explains the associations between discretionary accruals and beating the profit and earnings increase benchmarks.

One advantage of this study's analysis relative to performance-matched discretionary accruals (Kothari et al. 2005) is that it does not require specifying a metric to match upon (e.g., ROA)—i.e., it does not require specifying the exact nature of the correlation between discretionary accruals and performance. Relative to performance-matching this analysis may result in more powerful tests of earnings management around earnings benchmarks as it does not introduce "noise" in the tests via the matching process. Accordingly, this method may reduce the likelihood of falsely accepting the null hypothesis of no earnings management due to low-power tests.

This study's analyses should aid future researchers in investigating earnings management around earnings benchmarks using specific financial statement accounts (e.g., loan loss provision, valuation allowance, pension expense, etc.) or alternative abnormal accrual measures. In addition, this study's analyses should prove useful in tests for cross-temporal shifts in earnings management activity (e.g., pre- and post-Sarbanes-Oxley) or of other earnings management hypotheses (e.g., smoothing or big bath) that may be hindered by the systematic association between discretionary accrual proxies and performance. Finally, the *pseudo* target analyses may be particularly informative for studies assessing the properties of newly developed abnormal accrual measures.

This study's analyses should be interpreted in light of the following limitations. First, we assume that firms have greater incentives to manage earnings around the *actual* earnings benchmarks. If this assumption is not descriptive, then tests for increased associations between discretionary accruals and beating earnings benchmarks are not informative. Second, if firms manage earnings throughout the distributions of earnings, earnings changes, and unexpected earnings, then tests for increased associations between discretionary accruals and beating earnings benchmarks underestimate the association between discretionary accruals and beating the actual benchmarks that is attributable to earnings management. Third, our analyses assume that the underlying association between discretionary accruals and performance (i.e., irrespective of earnings management) does not shift around the *actual* earnings benchmarks. If this assumption is not descriptive, then tests could be biased in favor of or against finding more pronounced associations around the *actual* earnings benchmarks.

REFERENCES

Baber, W., and S-H. Kang. 2002. The impact of split adjusting and rounding on analysts' forecast error calculations. *Accounting Horizons* 16 (4): 277–289.

Beaver, W., M. McNichols, and K. Nelson. 2003. An alternative interpretation of the discontinuity in earnings distributions. Working paper, Stanford University.

Brown, L. D. 1991. Forecast selection when all forecasts are not equally recent. *International Journal of Forecasting* 7 (3): 349–356.

———, and K. Kim. 1991. Timely aggregate analyst forecast as better proxies for market earnings expectations. *Journal of Accounting Research* 29 (2): 382–385.

———, and M. Caylor. 2005. A temporal analysis of quarterly earnings thresholds: Propensity and valuation consequences. *The Accounting Review* 80 (2): 423–440.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (1): 99–126.

Dechow, P., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, ———, and ———. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13 (1): 1–36.

———, S. Richardson, and A. Tuna. 2000. Are benchmark beaters doing anything wrong? Working paper, University of Michigan.

———, ———, and ———. 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* 8 (2–3): 355–384.

DeFond, M., and K. R. Subramanyam. 1998. Auditor changes and discretionary accruals. *Journal of Accounting and Economics* 25 (1): 35–67.

Degeorge, F., J. Patel, and J. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (1): 1–33.

Durtschi, C., and P. D. Easton. 2005. Earnings management? Alternative explanations for observed discontinuities in the frequency distribution of earnings, earnings changes, and analyst forecast errors. *Journal of Accounting and Research* 43 (4): 557–592.

Guay, W., S. P. Kothari, and R. Watts. 1996. A market-based evaluation of discretionary-accrual models. *Journal of Accounting and Research* 34 (Supplement): 83–105.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (1–3): 85–107.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Kasznik, R. 1999. On the association between voluntary disclosure and earnings management. *Journal of Accounting Research* 37 (1): 57–81.

Kothari, S. P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (3): 483–514.

McNichols, M., and P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (Supplement): 1–31.

———. 2000. Research design issues in earnings management studies. *Journal of Accounting and Public Policy* 19 (4–5): 313–345.

O'Brien, P. C. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* 10 (1): 53–83.

Payne, J., and S. Robb. 2000. Earnings management: The effect of ex ante earnings expectations. *Journal of Accounting, Auditing, and Finance* 15 (4): 371–392.

———, and W. Thomas. 2003. The implications of using stock-split adjusted I/B/E/S data in empirical research. *The Accounting Review* 78 (4): 1049–1067.

Phillips, J., M. Pincus, and S. Rego. 2003. Earnings management: New evidence based on deferred tax expense. *The Accounting Review* 78 (2): 491–521.

Richardson, S., S. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. *Contemporary Accounting Research* 21 (4): 885–924.

Roychowdhury, S. 2003. Management of earnings through the manipulations of real activities that affect cash flow from operations. Working paper, Massachusetts Institute of Technology.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 653–676

Audit Partner Tenure and Audit Quality

Peter Carey
Monash University

Roger Simnett
The University of New South Wales

ABSTRACT: Rotation of audit partners is one of the main policy initiatives that has been implemented in many jurisdictions around the world to deal with concerns about audit quality. The basis of any requirement limiting the tenure of audit partners is that there is a reduction in audit quality associated with long periods of tenure. Using data from Australia, where the audit partner can be identified and for a period where partner rotation was not mandatory, we examine the association between audit quality and long audit partner tenure. The three measures of audit quality examined are the auditor's propensity to issue a going-concern audit opinion for distressed companies, the direction and amount of abnormal working capital accruals, and just beating (missing) earnings benchmarks. For long tenure observations we find a lower propensity to issue a going-concern opinion and some evidence of just beating (missing) earnings benchmarks, consistent with deterioration in audit quality associated with long audit partner tenure. There is no evidence of an association of long audit tenure with abnormal working capital accruals.

Keywords: *audit partner tenure; audit quality; qualifications; earnings management.*

Data Availability: *All data has been extracted from publicly available sources.*

I. INTRODUCTION

A policy of mandatory rotation of audit partners for a particular audit client is suggested as a means of improving audit quality and therefore increasing the quality of general-purpose financial statements. The benefits normally espoused are that the independence of the engagement partner is maintained or that a new perspective on the audit may result in the identification of issues that have been overlooked on previous audits. These benefits therefore rely on the belief that extended audit partner tenure is associated

---

We appreciate comments received on earlier versions of this paper from participants at the Accounting Association of Australia and New Zealand Conference, Boston Area Research Colloquium (BARC), University of Alberta, University of Connecticut, International Symposium on Audit Research University of Massachusetts (Amherst), University of Southern California, the University of Richmond, and the University of Antwerp. The comments of Jean Bedard, Stan Biggs, Jeff Cohen, Mike Gibbins, Ted Mock, Ken Trotman, Arnie Wright, Sally Wright, and anonymous reviewers are also much appreciated. We further appreciate the research assistance of Fernanda Afonso, Ana de Matos, Kim Huynh, Amna Khalifa, Bibi Moore, and Linda Pellegrino. Financial assistance was provided by a Monash University and University of New South Wales research grant.

Editor's note: This paper was accepted by Terry Shevlin.

Submitted January 2004
Accepted November 2005

with decreases in audit quality. This research study provides evidence on the extent to which long audit partner tenure is associated with decreases in audit quality.

In countries (such as the United States) where a policy of partner rotation is already in place, it is impossible to undertake an empirical evaluation of whether this policy is beneficial, as it is impossible to identify the consequences of not having (and therefore the benefits associated with implementing) this policy. The basis for the policy, that extended audit partner tenure is associated with decreased audit quality, is best informed by undertaking the analysis in an environment and at a time where the policy is not mandatory and the audit partner can be identified. Australia provides such an environment. There has been a statutory requirement for the auditor of listed Australian companies to sign the audit report in their own name as well as in the name of the audit firm since the 1970s (section 324 (10) of Australian Corporations Act). In Australia there was neither a legal nor professional requirement to rotate the audit partner until 2001.[1]

The basis for a policy of audit partner rotation is that there is a diminution in audit quality over the tenure of an audit partner. Such a policy assumes that a sufficiently high level of audit quality is achieved in the years up to recommended rotation. It is inferred that after a long period of tenure, continuing relationships may impact a partner's independence and time may erode a partner's capacity for critical appraisal. If these potential factors result in a reduction in audit quality or in audit failure, then they will impose a cost on society and the profession.[2]

This study investigates if there is a diminution in audit quality associated with long partner tenure using three common measures of audit quality: the auditor's propensity to issue a going-concern audit opinion; an examination of the signed and absolute amount of abnormal working capital accruals; and an analysis of the extent to which key earnings targets are just beaten (missed).

We undertake a cross-sectional analysis where we trace the length of partner tenure for all listed companies in Australia for a specific year to identify any association with audit quality. The year we choose is 1995, before a rotation policy became mandatory (2001) and before it was voluntarily adopted by any accounting firms (1997, refer to footnote 1). This time period allows for the identification of long tenure situations and therefore facilitates an examination of the theoretical underpinnings of an audit partner rotation policy; that there is a diminution in audit quality associated with long audit partner tenure.

For long partner tenure observations we find that there is a lower propensity to issue a going-concern opinion. There is no evidence of an association of long partner tenure with either the signed or absolute amount of abnormal working capital accruals. For the third audit-quality measure we find that there is some evidence of just beating (missing) earnings benchmarks for long partner tenure observations. The findings with regard the auditor's propensity to issue a going-concern audit opinion and the extent to which key earnings targets are just beaten (missed) are consistent with a deterioration in audit quality associated with long partner tenure. Sensitivity analyses demonstrate that identified deterioration in audit quality is specifically associated with non-Big 6 audit firms.

---

[1] Discussions with each of the Big 6 firms in Australia reveal that none of these firms had a partner rotation policy before 1997. Between 1997 and 2001 Arthur Andersen had a policy of mandatory partner rotation after seven years. The other Big 6 firms were aware of the issue, and undertook partner rotation on a best practice but not mandatory basis. In 2001, the accounting bodies in Australia introduced a professional requirement to rotate the audit partner for all listed companies at least every seven years.

[2] The auditing profession will collectively suffer loss with each audit failure through a reduction in the profession's credibility (Grant et al. 1996).

The paper is structured as follows. Section II outlines recent key policy initiatives regarding audit partner rotation in some of the major jurisdictions around the world. Section III outlines the literature review and development of theoretical arguments. This is followed by separate sections outlining the research methodology and the results for each of the three audit-quality measures. Section X concludes and outlines the implications of the findings.

II. RECENT POLICY DEVELOPMENTS REGARDING AUDIT PARTNER ROTATION IN MAJOR JURISDICTIONS

There is considerable variation in international policy regarding periodic audit partner rotation. The issue continues to attract policy attention in many countries. The policy first emerged in the U.S. in the 1970s, at which time the AICPA Practice Section mandated periodic audit partner rotation after seven years of tenure (this became a requirement for U.S. Securities and Exchange Commission registered clients). Recent corporate events led to a revision of this policy with the signing of the Sarbanes-Oxley Act of 2002 and there is now a requirement that the lead audit partner and audit review partner be rotated every five years on public company engagements.

In the United Kingdom, the Cadbury Committee (Committee on the Financial Aspects of Corporate Governance 1992) argued in favor of partner rotation and regulations were subsequently adopted that followed the U.S. approach of requiring rotation of the audit engagement partner after seven years for listed companies. In January 2003, the Department of Trade and Industry and the U.K. Treasury issued a final report on review of audit and accountancy regulations. The report recommended that the maximum period before rotation be reduced from seven to five years for the lead audit partner. The report also welcomed the recommendation from the EU that the partner rotation requirement should extend to other key audit partners (such as those involved in second partner review). It recommended a maximum seven-year period before rotation of other key audit partners.

More than 80 countries use the international auditing standards and Code of Ethics prepared by the International Federation of Accountants (IFAC) as the basis for their standards. In 2001 IFAC recognized in their Code of Ethics that prolonged use of the same lead engagement partner on an audit may create what they termed a "familiarity threat." The code proposed that the lead engagement partner be rotated after a pre-defined period, normally no more than seven years, and that the rotating partner not resume the lead engagement partner role until an additional two years have elapsed.

In Australia, the earliest references to partner rotation arose in a report issued by the professional accounting bodies (Australian Society of CPAs and The Institute of Chartered Accountants in Australia 1994). On the basis of concerns about audit quality they advocated, among other reforms, the rotation of audit engagement partners every seven years. This suggestion was not adopted by the profession. In 1996 an Australian Commonwealth Government working party released a review of the requirements for the regulation of auditors, which recommended mandatory rotation of audit partners every seven years. Again this suggestion was not adopted, although the independence standard at the time, AUP32, was revised to require "consideration" of a rotation policy. In 2001 the Australian profession's independence standard was revised in line with changes to the IFAC Code of Ethics. It stated in paragraph 2.51 that the lead partner be rotated after a predefined period of no longer than seven years. In 2004, legislation was introduced that requires the audit partner of listed companies to be rotated after no more than five years. There was no empirical evidence provided to support the benefits of either the introduction of a policy of partner rotation or a change in the conditions of this policy. As outlined earlier, the basis of any policy of audit partner rotation is that there is a diminution in the quality of the audit associated with long audit partner tenure. This study examines the existence of such an association.

III. LITERATURE REVIEW AND DEVELOPMENT OF TESTABLE HYPOTHESIS

While there are a number of recent studies on audit firm tenure (Ghosh and Moon 2005; Myers et al. 2003; Davis et al. 2003; U.S. Government Accounting Office [GAO] 2003; Geiger and Raghunandan 2002; Johnson et al. 2002), there is very little empirical examination of the relationship between audit partner tenure and audit quality. The only published paper to date to examine this issue is Chi and Huang (2005), who, in an examination of the Taiwanese audit market, find evidence consistent with lower earnings quality in the early years of audit tenure (either firm or partner) and the later years of tenure (firm), where earnings quality is measured as the level of abnormal accruals. However, their results are contingent on the method used to calculate the level of abnormal accruals.

The literature on audit firm tenure suggests deterioration in audit quality in the early years of tenure due to a loss of client-specific knowledge and expertise. DeAngelo (1981) identifies a "learning curve" that gives incumbent auditors a comparative quality advantage. Continuity on an audit is said to reduce audit risk due to a familiarity with the client's system and an understanding of risks associated with the client's business/industry environment (Financial Reporting Commission [Ryan Commission Report] 1992; AICPA [Cohen Commission Report] 1978).

While there is some empirical evidence of a heightened risk of reduced audit quality during the first years of an audit firm's tenure (AICPA 1992; Johnson et al. 2002; St. Pierre and Anderson 1984) and higher audit quality with longer audit firm tenure (Myers et al. 2003), the extent to which these relationships translate to audit partner tenure is not clear. The risk of audit failure in the years immediately after partner rotation is likely to be less severe than the risk in the first years of an audit firm's tenure due to other quality controls that an audit firm can establish to mitigate risks associated with a new partner's lack of familiarity (e.g., second partner review and standard new engagement familiarization procedures for incoming partner) and the maintenance of knowledge and expertise within the firm (e.g., continuity of field staff, the carrying forward of working papers, partner familiarity with existing audit methodology and client databases).

Association between Long Audit Partner Tenure and Audit Quality

The two (related) primary arguments supporting a negative association between long audit partner tenure and audit quality are (1) erosion of independence that may arise with the development of personal relationships between an auditor and their client, and (2) deterioration in the audit partner's capacity to effect critical appraisal (described in the IFAC Code of Ethics as a familiarity threat).

The development of personal relationships between client management and the audit partner may threaten the independence of the auditor. Mautz and Sharaf (1961, 231) state that the auditor "must be aware of the various pressures, some obvious some subtle, which tend to influence [their] attitude and thereby erode slowly but surely [their] independence." In most instances "the greatest threat to [their] independence is a slow, gradual, almost casual erosion of [their] honest disinterestedness" (Mautz and Sharaf 1961, 208). This threat to independence does not only occur at the time of reporting, but also has the potential to impact judgments made during the audit (Dopuch et al. 2003; Bazerman et al. 1997). Thus, it is alleged that extended personal relationships to the extent of developing bonds of loyalty or emotive relationships will consciously or subconsciously impact the auditor's independence and objectivity. Specific arguments as to the consequence of diminished independence with long partner tenure include the increased likelihood of the auditor yielding to the inevitable client pressure in an audit conflict situation and, at the extreme, the possibility that excessive familiarity might result in auditor/client collusion (McLaren 1958).

It is also argued that long audit partner tenure can lead to a familiarity threat that can result in a restriction on the creative audit-testing approaches that are often seen in the early years of an audit engagement (AICPA 1978; Hoyle 1978; McLaren 1958). The audit program may become stale and routine as the auditor begins to anticipate the condition of the client's systems and the presence of control procedures (Hoyle 1978). Shockley (1981, 789) asserts "complacency, lack of innovation, less rigorous audit procedures and a developed confidence in the client may arise after a long association with the client." There is a tendency to anticipate results instead of being alert to subtle and often surreptitious, although important, anomalies (Arrunada and Paz-Ares 1997).

There are also counter-arguments that suggest that long audit partner association with a particular client can result in higher audit quality. These arguments highlight the higher audit costs associated with early periods of auditor tenure and the increase in client and industry knowledge gained over repeated audits (Myers et al. 2003).

The increased prevalence of mandatory partner rotation policies shows that the regulators believe that the potential costs associated with long periods of partner tenure outweigh the potential benefits. The Cadbury Committee (1992) argues that partner rotation enhances the technical rigor of an audit by encouraging a "fresh viewpoint." The AICPA's early requirement to change the partner in charge of auditing a public company at least every seven years "was adopted for the specific purpose of periodically bringing a fresh perspective to each audit" (AICPA 1992, 4). The IFAC Code of Ethics (2003, para. 8.151) outlines that "using the same lead engagement partner on an audit over a prolonged period may create a familiarity threat." Partner rotation is therefore prescribed as a quality control procedure that overcomes the problems associated with long periods of partner tenure.

The prior discussion of a potential association between audit partner's tenure and audit quality suggests that in the early year(s) of tenure the partner is most likely to be independent and bring to the audit a fresh perspective. While the incoming engagement partner is developing the necessary knowledge and expertise, there is a possible increased risk of failure to detect material misstatements. The audit firm will attempt to reduce this risk through quality control processes designed to ensure the new engagement partner gains (or has access to) adequate knowledge and expertise. After this initial period has passed and the partner has gained the necessary client or industry knowledge and expertise, there is a period where audit quality is expected to be at its maximum. The relatively small number of years of association is not expected to have eroded independence or the partner's capacity to undertake critical appraisal. The length of this period is unknown, but the recent developments in policy initiatives outlined earlier in this study suggest that this period is commonly viewed as either five or seven years. After this period, the literature suggests a reduction in audit quality associated with long audit partner tenure through the development of personal relationships and deteriorating capacity for critical appraisal (familiarity threat). This association is examined by the following hypothesis:

H1: There is a negative association between audit quality and long audit partner tenure.

In the analyses undertaken in the paper, long audit partner tenure is defined as situations where the auditor has been a partner in charge for more than seven years (*TENURE>7*). This is consistent with the period of concern currently espoused by the IFAC Code of Ethics, and the period of partner tenure that was commonly used in 1995 in other jurisdictions as the basis for a partner rotation policy. We also anticipate that there may be an initial period of familiarization for the audit partner. Thus we include the first two years of partner tenure as a control variable (*TENURE≤2*). This approach means that the basis of comparison for the later partner tenure is the period where audit quality is expected to be at its maximum (for this paper, between years 3 to 7 of partner tenure). Other possible specifications of long partner tenure are considered as forms of sensitivity analyses where appropriate. This includes the use of a continuous measure for partner tenure, *TENURE>5*, and to further accentuate the effect of long periods of tenure, tenure squared.

IV. METHODOLOGY

Data Collection

The data collection for this study involves the review and analysis of published information for public companies listed on the Australian Stock Exchange (ASX) in 1995. We select 1995 as the year for analysis as this is before the introduction of firm, professional, or legal requirements to limit audit partner tenure in Australia. For all listed companies in 1995 we trace back the partner tenure, as well as collecting other relevant information for the analyses of the audit-quality measures. The population of companies listed on the ASX at 30 June 1995 is 1,152 (Craswell 1996). Of these, 40 foreign companies are removed (the audit report is signed in an overseas jurisdiction and does not necessarily identify the audit partner) and 66 are removed due to missing data with respect to incomplete financial information or an inability to trace the inception of audit partner tenure. The 25 observations in the banking and insurance industry are also excluded as their total asset base and financial structure is not comparable to those of the other companies. This results in a final sample of 1,021 Australian-domicile companies.

Research Method

To identify whether there is a negative association between long audit partner tenure and audit quality, we use three common measures of audit quality.[3] First, the auditor's propensity to issue a going-concern opinion for distressed companies; second, the amount of abnormal working capital accruals; and third, the extent to which key earnings targets are just beaten (missed). These are all examined over short, medium, and long periods of audit partner tenure. In accordance with the predicted directional effect, we utilize one-tailed testing in evaluating the impact of audit partner tenure. The comparison of measures of audit quality for the medium (when audit quality is expected to be at its maximum) with the long tenure periods allows us to identify any diminution in audit quality associated with long tenure. The examination of multiple measures of audit quality allows us to build greater confidence in any observed relationship between audit quality and long partner tenure. The research methods associated with each of these measures are discussed with their results.

---

[3] Other measures considered are the number of audit failures as identified by legal action undertaken against the auditor, or companies placed in bankruptcy soon after the issue of an unmodified audit report. The very small number of legal actions and bankruptcies in Australia means that these proxies are not appropriate.

V. DESCRIPTIVE RESULTS AND DISCUSSION

Descriptive statistics are reported in Table 1. Of the 1,021 Australian-listed companies, 15 percent (n = 152) have greater than seven years of continuous partner tenure.[4] Forty-four percent of companies (n = 447) have less than two years of continuous partner tenure. The companies have been listed on average 12.81 years and have A$299 million of assets. The Big 6 audit firms audit 64 percent of these clients and 7 percent of the audit opinions are modified relating to the going-concern issue. An examination of the correlation matrix reveals that apart from the expected high correlations of *PBANK* with *LEV* (0.627) and *CFFO* (−0.532), there are no other correlations greater than 0.50.

VI. ANALYSIS 1: ISSUING GOING-CONCERN AUDIT OPINIONS

Method

The first measure of audit quality examined in this study is that of the auditor's propensity to issue a going-concern modified audit opinion. Although it is not easy to associate

TABLE 1
Descriptive Statistics
n = 1,021

Variables	Mean	Median	Std. Deviation	Minimum	Maximum
TENURE>7	0.15	0	0.36	0	1
TENURE≤2	0.44	0	0.50	0	1
PBANK	−1.64	−2.36	4.08	−7.89	33.57
OPINION	0.07	0	0.25	0	1
SIZE ($ in millions)	299.08	22.80	1,642.28	.01	30,287.00
AGE	12.81	9.00	12.94	1	93
LEV	0.36	0.32	0.34	.00	2.15
LLOSS	0.38	.00	0.49	0	1
AUDFIRM	0.64	1.00	0.480	0	1
FEERATIO	0.33	0.31	0.25	.00	.95
CFFO	−0.03	0.02	0.24	−1.62	0.51
MINING	0.37	0	0.48	0	1

Variable Definitions:

TENURE>7 = 1 if audit partner is engagement partner on a client company more than seven years, and 0 otherwise;
TENURE≤2 = 1 if audit partner is engagement partner on a client company less than or equal to two years, and 0 otherwise;
OPINION = 1 if company receives a going-concern modified opinion, and 0 otherwise;
PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;
SIZE = total assets of the company at financial year-end;
AGE = number of years since the company was listed on the Australian Stock Exchange;
LEV = total liabilities divided by total assets;
LLOSS = 1 if company reported a loss for the previous year, and 0 otherwise;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
FEERATIO = ratio of nonaudit fees to total fees paid to the incumbent auditor;
CFFO = cash flow from operations over total assets at the end of the fiscal year; and
MINING = 1 if mining industry, and 0 otherwise.

---

[4] Twenty-eight percent (n = 286) have more than five years of continuous partner tenure. The average audit partner tenure is just over four years, ranging from 1–24 years.

any particular type of audit opinion with a level of audit quality, there is a well-established body of literature that aids the prediction of going-concern modified opinions (see for example, DeFond et al. 2002; Mutchler et al. 1997; Hopwood et al. 1989). DeFond et al. (2002) argue that in issuing a going-concern opinion, the auditor must objectively evaluate company performance and withstand client pressure to issue a clean opinion. This suggests a correlation between the issuing of going-concern opinions and auditor independence. Thus, if after controlling for other factors that influence the propensity to issue a going-concern modified audit opinion, the probability of issuing going-concern reports is inversely related to long audit partner tenure, this suggests audit quality is diminishing with long partner tenure.

Consistent with DeFond et al. (2002) and on the basis that going-concern problems are more salient among financially distressed companies, we restrict our analysis to situations where the company reports negative earnings and/or operating cash flows for the current fiscal year. Also, consistent with the approach of DeFond et al. (2002), the following logistic regression model estimates the auditor's probability of issuing a going-concern modified opinion to a financially distressed client:

$$\begin{aligned} OPINION = {} & \beta_0 + \beta_1 PBANK + \beta_2 SIZE + \beta_3 AGE + \beta_4 LEV + \beta_5 CLEV \\ & + \beta_6 RETURN + \beta_7 LLOSS + \beta_8 INVESTMENTS + \beta_9 AUDFIRM \\ & + \beta_{10} FEERATIO + \beta_{11} CFFO + \beta_{12} MINING + \beta_{13} TENURE{\leq}2 \\ & + \beta_{14} TENURE{>}7 + \varepsilon \end{aligned} \tag{1}$$

where:

Dependent Variable:

OPINION = 1 if company receives a going-concern modified opinion, and 0 otherwise;

Experimental Variable:

TENURE>7 = 1 if audit partner is engagement partner on a client company for greater than seven years, and 0 otherwise;

Control Variables:

PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;[5]
SIZE = natural logarithm of total assets of the company at financial year-end;
AGE = natural logarithm of number of years since listing on the Australian Stock Exchange;
LEV = total liabilities divided by total assets;
CLEV = change in *LEV* during the year;
RETURN = market-adjusted return over the fiscal year;

---

[5] This measure is used in many studies (e.g., Carcello et al. 1995) and incorporates the financial variables found to be significant in bankruptcy prediction studies (e.g., Hopwood et al. 1989). This is based on the form of the model calculated by Carcello et al. (1995) who use the weightings $b^* = -4.803 - 3.6$(net income/total assets) $+ 5.4$(total debt/total assets) $- 0.1$(current assets/current liabilities). There are no significant differences in the results when the original form of the bankruptcy model as developed by Zmijewski (1984) is used.

LLOSS = 1 if client reported a loss for the previous year, and 0 otherwise;
INVESTMENTS = short- and long-term investment securities (measured as current assets less debtors and inventory) divided by total assets;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
FEERATIO = ratio of nonaudit fees to total fees paid to the incumbent auditor;
CFFO = operating cash flow divided by total assets;
MINING = 1 if mining industry, and 0 otherwise; and
TENURE≤2 = 1 if audit partner is engagement partner on a client company for two years or less, and 0 otherwise.

The choice of control variables in the model is informed by the literature that aids the prediction of a going-concern opinion (in particular, DeFond et al. 2002). As outlined earlier, *TENURE≤2* is used to control for instances of possible lower audit quality in the first two years of partner tenure due to a lack of client familiarity affecting auditor expertise. The other variables are consistent with the model of DeFond et al. (2002).[6] *PBANK* is the probability of bankruptcy score based on Zmijewski (1984), with higher values indicating a higher probability of bankruptcy. *SIZE* is included because large companies have greater negotiating power and are less likely to end up in bankruptcy. *AGE* captures the fact that younger companies are more likely to encounter financial distress. *LEV* captures risk associated with higher levels of debt, and *CLEV* captures movements in leverage that may move companies to unacceptable or unsustainable levels of debt. *RETURN* is a market-based measure of risk and company performance. *LLOSS* is included as companies with continued losses are more likely to fail. *INVESTMENTS* is a liquidity measure that captures the ability to quickly raise cash, while *AUDFIRM* captures any differences in the propensity of the big audit firms to issue going-concern modifications to the audit opinion. It is suggested that *FEERATIO* is associated with reduced audit quality (Frankel et al. 2002) and poor operating cash flows (*CFFO*) are commonly associated with a higher likelihood of financial distress (DeFond et al. 2002). The variable *MINING* is included because of the large number of mining companies listed in Australia and their potentially different financial profiles (Butterworth and Houghton 1995).

Results

The restriction of our analysis to situations where the client reports negative earnings and/or operating cash flows in 1995 results in the identification of 559 companies. The average partner tenure of these observations is 3.75 years (46 percent are in either years 1 or 2 of partner tenure and 11 percent are long tenure observations of greater than seven years) and we observe 66 (12 percent) going-concern modified audit opinions.

Table 2 presents results from the logistic regression model for going-concern modified audit opinions. The model adequately distinguishes the going concern decision (χ^2 = 114.15, df = 14, p = .000), and an overall fit of 88.6 percent is achieved. Consistent with DeFond et al. (2002), we find significance in the predicted direction at $p < .05$ for the coefficients *LLOSS* and *INVESTMENTS*, while *SIZE, CLEV, RETURN, CFFO,* and *FEERATIO* are not found to be significant predictors. However, in contrast to DeFond et al. (2002), we find significance in the predicted direction for *LEV* and opposite to the predicted direction for *AGE*. We do not find significance in the predicted direction of the

---

[6] Variables used by DeFond et al. (2002) but not included in the present study are the market variables of beta risk and return volatility and an indicator variable for ability to raise future financing. These variables are not available for many of our observations and are therefore not considered.

TABLE 2

Propensity to Issue a Going-Concern Opinion for Financially Distressed Companies

(n = 559)

$$\begin{aligned} OPINION = {} & \beta_0 + \beta_1 PBANK + \beta_2 SIZE + \beta_3 AGE + \beta_4 LEV + \beta_5 CLEV + \beta_6 RETURN \\ & + \beta_7 LLOSS + \beta_8 INVESTMENTS + \beta_9 AUDFIRM + \beta_{10} FEERATIO \\ & + \beta_{11} CFFO + \beta_{12} MINING + \beta_{13} TENURE{\leq}2 + \beta_{14} TENURE{>}7 + \varepsilon. \end{aligned}$$

	Expected Sign	**Coefficient (p-value)**
Experimental Variable		
TENURE>7	−	−1.517 (.042)**
Control Variables		
TENURE≤2	−	.487 (.072)*
PBANK	+	.035 (.325)
SIZE	−	−.061 (.591)
AGE	−	.453 (.022)**
LEV	+	1.809 (.000)**
CLEV	+	−.477 (.337)
RETURN	−	−.243 (.295)
LLOSS	+	1.748 (.000)**
INVESTMENTS	−	−2.411 (.005)**
AUDFIRM	+	−.630 (.053)*
FEERATIO	−	−.362 (.611)
CFFO	−	.187 (.790)
MINING	?	−.405 (.236)
Constant		−3.063 (.108)

Pseudo $R^2 = 0.358$

Model Chi-square = 114.15, $p < 0.00$

---

*, ** Significant at the 10 percent and 5 percent levels, respectively.

Variable Definitions:

OPINION = 1 if company receives a going-concern modified opinion, and 0 otherwise;
TENURE>7 = 1 if audit partner is engagement partner on a client company more than seven years, and 0 otherwise;
TENURE≤2 = 1 if audit partner is engagement partner on a client company less than or equal to two years, and 0 otherwise;
PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;
SIZE = the natural logarithm of total assets of the company at financial year-end;
AGE = the natural logarithm of the number of years since the company was listed;
LEV = total liabilities divided by total assets;
CLEV = change in *LEV* during the year;
RETURN = market-adjusted return over the fiscal year;
LLOSS = 1 if company reported a loss for the previous year, and 0 otherwise;
INVESTMENTS = short- and long-term investment securities (measured as current assets less debtors and inventory) divided by total assets;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
FEERATIO = ratio of nonaudit fees to total fees paid to the incumbent auditor;
CFFO = cash flow from operations over total assets at the end of the fiscal year; and
MINING = 1 if mining industry, and 0 otherwise.

coefficients for *AUDFIRM* and *PBANK*. The estimated coefficient for *TENURE≤2* is positive but is also not significant ($p = .072$, one-tailed).

The negative estimated coefficient for *TENURE>7* is significant ($p = .042$, one-tailed) and in the expected direction. This result suggests that long-serving audit partners have a lower propensity to issue a going-concern modified opinion. This provides evidence that is consistent with H1, that long audit partner tenure is negatively associated with audit quality. To determine the economic significance of this result, we change the value for *TENURE>7* from 0 to 1 while setting other variables in model (1) equal to their sample mean. This results in a 5.30 percent decrease in the probability of receiving a going-concern modified opinion.

Sensitivity Analyses

To control for the impact of a prior year going-concern opinion on an auditor's propensity to issue a going-concern opinion in the current year, we first replicate the primary logistic regression analysis, substituting first-time going-concern opinions as the dependent variable, as in DeFond et al. (2002). Of the 66 going-concern modifications. there are 33 first-time going-concern opinions.[7] The model is a significant predictor of first-time going-concern modifications ($\chi^2 = 49.54$, df = 14, p = .000) and results are broadly consistent with the analysis containing all going-concern modifications. The *TENURE>7* variable has a negative coefficient but its significance is now marginal (p = .075, one-tailed). Second, we rerun the primary analysis including a dummy variable for whether the company receives a prior-year going-concern modification. Similar results to those of the analysis of the first-time going-concern opinions are obtained, with *TENURE>7* having a negative coefficient and again being marginally significant (p = .098, one-tailed).

Identifying that not all loss making or cash flow negative firms will receive a going-concern modified opinion (only 66 of the 559 companies received going-concern modifications), we extend this analysis to consider those companies most likely to receive a going-concern modified opinion. We therefore examine the most severely distressed quartile (140 companies) of the financially stressed companies.[8] The going-concern modified opinion is issued in 17 of 43 (39.5 percent) instances for short tenure, 27 of 80 (33.8 percent) instances for medium tenure, and 2 of 17 (11.8 percent) instances for long tenure (greater than seven years). Using a test of proportions, we identify that there is no significant difference between the proportions for short and medium tenure (z = 0.64, p > .05), while the difference between medium and long tenure is significant (z = 1.72, p = .043, one-tailed). This provides evidence of a lower propensity to modify the audit opinion for the most severely financially stressed companies associated with long audit partner tenure.

To further investigate if the auditor's propensity to issue a modified audit report diminishes with tenure, we substitute the two tenure variables (*TENURE≤2*, *TENURE>7*) with a single linear *TENURE* variable, while controlling for instances of possible lower audit quality in the first two years of partner tenure by excluding the 258 *TENURE≤2* observations. *TENURE* remains negative and significant (p = .012, one-tailed),[9] indicative of a lower propensity to issue a going-concern modification during long partner tenure.

To explore the potential interaction between *AUDFIRM* and long partner tenure, we include an interaction term *AUDFIRM*TENURE* in model (1). *TENURE* is measured as a continuous variable[10] and observations where tenure is less than two years are excluded. The inclusion of the interaction term results in *TENURE* remaining significant and negative ($\beta = -.509$, p = .005) while the interaction term is significant and positive ($\beta = .544$, p = .021). This analysis shows that the significant finding for *TENURE* identified in the earlier analysis is due to non-Big 6 firms having a lower propensity to issue a going-concern modified opinion than Big 6 firms in the later years of partner tenure.[11]

---

[7] The 33 ongoing going-concern modification observations are excluded from the analysis.

[8] Those companies with the greatest possibility of bankruptcy are identified by applying model (1). This quartile received 46 of the 66 (69.7 percent) going-concern modified audit opinions.

[9] When we square the tenure variable to accentuate the effect of partner tenure in the full sample of 559 financially distressed companies, the coefficient remains negative and highly significant (p = .003, one-tailed).

[10] We use the continuous *TENURE* measure because the small number of going-concern modifications for *TENURE* > 7 observations results in a large error term and unreliable results.

VII. ANALYSIS 2: ABNORMAL WORKING CAPITAL ACCRUALS

Method

Consistent with prior literature, we posit that higher-quality audits mitigate extreme management reporting decisions. Accruals are commonly used to identify these extreme reporting decisions (Myers et al. 2003; Francis et al. 1999; Becker et al. 1998). We concentrate on abnormal working capital accruals (*AWCA*) and utilize the proxy measure of *AWCA* developed by DeFond and Park (2001). *AWCA* is the difference between realized working capital and an expected level of working capital needed to support a current sales level, where an historic relation of working capital to sales captures expected working capital. In their analysis, DeFond and Park (2001, 399) find their measure to be a more powerful test in comparison to a test using total (normal plus abnormal) accruals. The focus on working capital accruals is further supported by prior research that suggests that management has the most discretion over such accruals (Ashbaugh et al. 2003; Becker et al. 1998).

The measure of *AWCA* is:

$$AWCA_t = WC_t - [(WC_{t-1}/S_{t-1}) \times S_t]$$

where:

t = year, $t-1$ refers to the prior year;
WC_t = non-cash working capital in the current year computed as (current assets − cash and short-term investments) − (current liabilities − short-term debt);
WC_{t-1} = non-cash working capital in the prior year;
S_t = sales in current year; and
S_{t-1} = sales in prior year.

We scale all variables by average total assets (Myers et al. 2003).

We examine the signed and absolute amounts of *AWCA*, which are indicators of the degree to which management is allowed to exercise discretion in reporting earnings. If reduced audit quality occurs in the later years of audit partner tenure, then we expect larger (income-increasing) signed and higher unsigned (absolute) amounts of *AWCA* in these later years (Myers et al. 2003; Frankel et al. 2002; Reynolds and Francis 2000; Francis et al. 1999; Becker et al. 1998).

We estimate the following model:

---

[11] For the Big 6 sample the qualification rate is 8 of 133 observations (6.0 percent) for medium tenure compared with 2 of 34 observations (5.9 percent) for long tenure. For the non-Big 6 sample the qualification rate is 24 of 106 observations (22.6 percent) for medium tenure and 0 of 28 observations (0 percent) for long tenure.

$$\begin{aligned} AWCA = {} & \beta_0 + \beta_1 AUDFIRM + \beta_2 PBANK + \beta_3 OPINION + \beta_4 SIZE + \beta_5 LEV \\ & + \beta_6 LLOSS + \beta_7 FEERATIO + \beta_8 PERFORM + \beta_9 AGE \\ & + \beta_{10} GROWTH + \beta_{11} CFFO + \beta_{12} MINING + \beta_{13} TENURE{\leq}2 \\ & + \beta_{14} TENURE{>}7 + \varepsilon \end{aligned} \tag{2}$$

where:

Dependent Variable:

AWCA = abnormal working capital accruals (measured in signed or absolute amount);

Additional Control Variables:

PERFORM[12] = earnings before tax over total assets at the end of the fiscal year; and
GROWTH = change in assets from prior year.

The choice of control variables is informed by prior studies that examine aspects of audit quality by measuring the level of abnormal accruals, in particular those conducted by Myers et al. (2003), Johnson et al. (2002), and Frankel et al. (2002). A number of the variables (*PBANK, SIZE, LEV, LLOSS, AUDFIRM, OPINION, FEERATIO, AGE, CFFO, MINING*) contained in the going-concern model outlined in the previous section are expected to be associated with both going-concern prediction and the level of abnormal accruals. Two of the control variables in the going-concern model (*INVESTMENTS* and *CLEV*), are not used in abnormal accrual studies and, thus, are not included in this study. Two additional variables are included. As abnormal accrual models do not completely extract nondiscretionary accruals that are correlated with company performance (e.g., Dechow et al. 1995; Ashbaugh et al. 2003), we include a measure of performance to control for the nondiscretionary component of *AWCA* not extracted by the model (*PERFORM*). As accruals are likely to be associated with a company's growth opportunities (Johnson et al. 2002), *GROWTH* is controlled for using the change in assets from the prior year.

Results

The sample for which *AWCA* are calculated is reduced from 1,021 to 743 companies. Fifty-six companies are deleted because the required prior-year information is not available (newly listed companies), and 222 because the company does not report a consistent pattern of sales (revenue) between the current and previous fiscal years.[13]

Table 3 presents the results of estimating a model examining the relationship between *AWCA* (as an indicator of earnings management) and audit partner tenure. With respect to the control variables, *CFFO* is negatively associated with *AWCA* (consistent with Myers et al. [2003] and Johnson et al. [2002]). Riskier companies and more highly levered companies are also found to have higher levels of signed *AWCA*s. There is some support that signed

---

[12] When we substitute *PERFORM* with the absolute value of *PERFORM*, there is no significant differences in the results.

[13] Unstable sales revenue is determined as zero revenue in any year, or less than 10 percent of the comparison year's sale revenue.

TABLE 3
OLS Regression Results for Abnormal Working Capital Accruals
(n = 743)

$$AWCA = \beta_0 + \beta_1 AUDFIRM + \beta_2 PBANK + \beta_3 OPINION + \beta_4 SIZE + \beta_5 LEV + \beta_6 LLOSS + \beta_7 FEERATIO + \beta_8 PERFORM + \beta_9 AGE + \beta_{10} GROWTH + \beta_{11} CFFO + \beta_{12} MINING + \beta_{13} TENURE{\leq}2 + \beta_{14} TENURE{>}7 + \varepsilon.$$

		Signed *AWCA*		**Absolute *AWCA***	
Variables	**Expected Sign**	**Coefficient**	**t-value**	**Coefficient**	**t-value**
Experimental Variable					
TENURE>7	+	−.017	(−.247)	−.027	(−.454)
Control Variables					
TENURE≤2	+	.087	(1.648)*	.030	(.658)
PBANK	+	−.118	(−2.333)**	−.002	(−.049)
OPINION	+	.148	(1.207)	.112	(1.063)
SIZE	−	.012	(.853)	−.028	(−2.274)**
LEV	+	.604	(2.002)**	.129	(.497)
LLOSS	+	.047	(.757)	.144	(2.672)**
AUDFIRM	−	.072	(1.352)	.001	(.017)
FEERATIO	−	−.193	(−1.842)*	.108	(1.202)
PERFORM	?	−.304	(−1.113)	.173	(.735)
AGE	−	.031	(1.099)	.027	(1.109)
GROWTH	+	−.004	(−.323)	−.012	(−1.004)
CFFO	−	−.324	(−2.023)**	−.305	(−2.212)**
MINING	?	−.034	(−.595)	.036	.732
Constant		−.783	(−2.348)**	.481	(1.672)*
Adj. R^2		.012		.040	

*, ** Significant at the 10 percent and 5 percent levels, respectively.

Variable Definitions:

AWCA = abnormal working capital accruals (measured in signed or absolute amount);
TENURE>7 = 1 if audit partner is engagement partner on a client company more than seven years, and 0 otherwise;
TENURE≤2 = 1 if audit partner is engagement partner on a client company less than or equal to two years, and 0 otherwise;
PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;
OPINION = 1 if company receives a going-concern modified opinion, and 0 otherwise;
SIZE = the natural logarithm of total assets of the company at financial year-end;
LEV = total liabilities divided by total assets;
LLOSS = 1 if company reported a loss for the previous year, and 0 otherwise;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
FEERATIO = ratio of nonaudit fees to total fees paid to the incumbent auditor;
PERFORM = earnings before tax over total assets at the end of the fiscal year;
AGE = the natural logarithm of the number of years since the company was listed;
GROWTH = change in assets from prior year;
CFFO = cash flow from operations over total assets at the end of the fiscal year; and
MINING = 1 if mining industry, and 0 otherwise.

*AWCA*s are higher in the earlier years of partner tenure. Smaller companies and those with a prior year loss have higher absolute levels of *AWCA*.

If the level of earnings management increases as audit partner tenure increases, then we expect higher levels of signed and absolute accruals during the later years of partner tenure. For both the measures of *AWCA* (signed and absolute), the models are significant and the amount of explained variation ranges from 1.2 percent to 4.0 percent. For both models, *TENURE>7* is not significant, suggesting that earnings management as measured by *AWCA* is not associated with long partner tenure.

Because of asymmetric relationships between the independent variables and positive and negative abnormal accruals, income-increasing and income-decreasing abnormal accruals are analyzed separately (Menon and Williams 2004; Myers et al. 2003). Accordingly we separate the full sample into companies with positive (income-increasing) and negative (income-decreasing) *AWCA* and estimate the regression models separately on these two samples.[14] For the reduced sample of income-increasing *AWCA*s ($n = 398$), the explanatory power of the model improves (adjusted $R^2 = .048$) and *TENURE>7* is not significant ($t = -.069$, $p > .05$). For the sample of income-decreasing *AWCA*s ($n = 345$), the explanatory power of the model is lower (adjusted $R^2 = .022$) and *TENURE>7* is not significant ($t = -.150$, $p > .05$).[15] The results therefore suggest no deterioration in audit quality associated with long partner tenure, for either income-increasing or -decreasing *AWCA*.

Overall, our results do not find that long audit partner tenure is associated with either a deterioration in audit quality (as hypothesized), or an improvement in audit quality (as suggested by Myers et al. [2003] when examining audit firm tenure), when audit quality is measured by its association with the signed or absolute amounts of *AWCA*.

Sensitivity Analyses

The fact that *TENURE>7* is not significant may be caused by the small number of long tenure observations. Substituting the experimental variables with a binary variable measuring tenure greater than five years (*TENURE>5*) again finds no evidence that long partner tenure is associated with reduced audit quality.

As income-increasing *AWCA* are associated with lower audit quality (Menon and Williams 2004; Myers et al. 2003; DeFond and Park 2001; Becker et al. 1998; DeFond and Jiambalvo 1993), we also examine the sign (income-increasing versus income-decreasing) of *AWCA*. The logistic regression model is significant, and the amount of explained variation is 5.7 percent. Neither *TENURE>7* or *TENURE≤2* are significant. This suggests that earnings management as measured by the sign of *AWCA* is not associated with partner tenure.

To further investigate whether audit partner tenure is associated with *AWCA*, we replace the *TENURE≤2* and *TENURE>7* variables with a single linear *TENURE* variable and control for instances of possible lower audit quality in the first two years of partner tenure by excluding the 296 *TENURE≤2* observations. Again, no evidence is found to suggest that long periods of partner tenure are associated with reduced audit quality.[16]

In the analyses undertaken, *AUDFIRM* is not significant despite a good *a priori* reason to believe that long tenure plays out differently for larger audit firms (as previously discussed under sensitivity analysis for Section VI). To further explore the potential interaction between long partner tenure and *AUDFIRM*, we use the same approach as that used in Section VI. This involves excluding the first and second year of tenure observations and including *TENURE* as a continuous variable and *AUDFIRM*TENURE* as an interaction term in model (2). Both *TENURE* and the interaction term are insignificant. This suggests that there are no differences in the signed or absolute level of accruals over different levels of tenure for Big 6 versus non-Big 6 audit firms.

---

[14] As the samples are truncated at one end of the distribution for both regressions, we use a maximum likelihood truncated regression procedure to obtain corrected coefficients.

[15] *TENURE≤2* is not significant in either analysis.

[16] When we square the tenure variable to accentuate the effect of partner tenure, the coefficient remains nonsignificant.

As part of the sensitivity analysis, we also run a variation of the definition of *AWCA*, utilizing the measure of current accruals used by Myers et al. (2003).[17] This is:

$$Curr\_Accruals = ((\Delta Curr\_Ass - \Delta Cash) - (\Delta Curr\_Liab - \Delta STD))$$

where:

$\Delta Curr\_Ass$ = change in current assets;
$\Delta Cash$ = change in cash equivalents;
$\Delta Curr\_Liab$ = change in current liabilities; and
ΔSTD = change in short-term notes and current portion of long-term debt.

Using this measure of current accruals (either signed or absolute) as the dependent variable for model (2) does not change the insignificance of the *TENURE>7* variable.

VIII. ANALYSIS 3: JUST BEATING OR MISSING EARNINGS BENCHMARKS

Method

For this analysis, the quality of audited financial information is measured by the extent of earnings management. This is evidenced by whether the earnings just beats or misses specific benchmarks, and the extent to which this event is associated with auditor tenure. Benchmark studies find (1) companies avoid reporting a loss; (2) an unusually low frequency of small decreases in earnings and unusually high frequency of small increases in earnings;[18] and (3) an unusually high frequency of companies meeting or exceeding analysts' forecasts (Menon and Williams 2004; Holland and Ramsay 2003; Matsumoto 2002; Degeorge et al. 1999; Burgstahler and Dichev 1997; Hayn 1995). We utilize the first two measures[19] in order to identify the association between earnings management and partner tenure. The formulae are:

(a) Avoid reporting a loss = $Earnings_t / Total\ Assets_t$; and
(b) Exceed prior year's earnings = $(Earnings_t - Earnings_{t-1}) / Total\ Assets_{t-1}$.

We utilize the model used in Menon and Williams (2004) on meeting or missing benchmarks to investigate the relationship between audit partner tenure and audit quality. We therefore develop the following multivariate logistic regression models:

$$\begin{aligned} Y = {} & \beta_0 + \beta_1 AUDFIRM + \beta_2 PBANK + \beta_3 SIZE + \beta_4 AGE + \beta_5 MKTVAL \\ & + \beta_6 MINING + \beta_7 TENURE{\leq}2 + \beta_8 TENURE{>}7 + \varepsilon \end{aligned} \qquad (3)$$

---

[17] This measure has the advantage that it is possible to retain the 222 companies that do not report a consistent pattern of sales between the current and previous fiscal years. (The analysis reported here does not include these companies, but they are included for the sensitivity analysis contained in Section IX.) The measure however has the disadvantage of not relating these accruals to a level of activity for the company, and therefore fails to identify abnormal working capital accruals for that company.

[18] DeAngelo (1986) finds evidence of strong market incentives to maintain consistent earnings increases, and Barth et al. (1999) observe that managers focus on maintaining steadily increasing earnings. In a field study, Nelson et al. (2002) also provide evidence that managers are more likely to make attempts that increase current year's income.

[19] The availability of analysts' forecasts is limited to only the very large companies in Australia in 1995, resulting in insufficient degrees of freedom to undertake this analysis.

where Y is defined in four different models as $\Pr(BEATS\_BE = 1)$, $\Pr(MISSES\_BE = 1)$, $\Pr(BEATS\_LYR = 1)$, and $\Pr(MISSES\_LYR = 1)$.

Additional Control Variable:

MKTVAL = the natural logarithm of the market value of equity.

Using a similar method of analysis to Menon and Williams (2004), there are two regression analyses for the first benchmark (company performance around breakeven). In the first regression analysis the dependent variable *BEATS_BE* = 1 when profit is less than 2 percent of total assets, and 0 otherwise. In the second regression analysis the dependent variable *MISSES_BE* = 1 when loss is less than 2 percent of total assets, and 0 otherwise. For the second benchmark (last year's profit), the dependent variable *BEATS_LYR* = 1 when the increase in profit (decrease in loss) over last year's profit/loss is less than 2 percent of total assets, and 0 otherwise. We then undertake the analysis where the dependent variable *MISSES_LYR* = 1 when decrease in profit (increase in loss) over last year's profit/loss is less than 2 percent of total assets, and 0 otherwise. If clients with long tenure partners are granted more discretion than other clients, then we expect β_8 to be positive and significant for the analyses of the just beats the benchmarks (breakeven and last year's profit) and negative and significant for the analyses of the just misses the benchmarks.

To determine whether the just beats or misses specific earnings benchmarks are associated with long partner tenure (*TENURE>7*), we control for the age of the client (*AGE*) and variations in the size of clients by using both the natural logarithm of total assets (*SIZE*) and market value of equity (*MKTVAL*) (as in Menon and Williams 2004). In addition to *TENURE≤2* we include three additional variables not in Menon and Williams (2004). As previously discussed, audit firm type is expected to constrain earnings management and is controlled for by including the variable *AUDFIRM*. *PBANK*, which is measured by an adjusted Zmijewski score, is also expected to be associated with earnings management. To control for the high proportion of mining companies in Australia we include the variable *MINING*. The analysis is performed on the population of 1,021 company observations.

Table 4, Panel A shows the frequency with which short, medium, and long partner tenure are associated with just beats versus just misses breakeven. For short partner tenure, the proportion of the combined just beats and just misses observations that do just beat breakeven is 26 of 43 (60.5 percent). For medium tenure, this is 22 of 45 (48.9 percent) and for long tenure, this is 11 of 15 (73.3 percent). The proportion of long tenure versus medium tenure observations that just beats breakeven versus just misses breakeven is significant ($z = 1.65$, $p = .05$, one-tailed). This is consistent with the view that there is a lower audit quality associated with long audit partner tenure.

Separate regression analyses are calculated for the two dependent variables; just misses breakeven and just beats breakeven. The results of these are contained in Table 4, Panel B. Consistent with the descriptive results above, *TENURE>7* is positive but not significant for the just beats breakeven analysis, and negative and approaching significance for the just misses breakeven. (*TENURE≤2* is also negative and approaching significance for the just misses breakeven.)

In the second analysis, which examines the just beats (just misses) last year's profit (by less than 2 percent of total assets), the 56 first-year listing observations are deleted, leaving 965 observations in the analysis. Table 5, Panel A shows the frequency with which short, medium, and long partner tenure are associated with just beats versus just misses

TABLE 4
Just Beats and Just Misses Breakeven
(n = 1,021)

Panel A: Earnings Benchmark Frequencies and Statistics

	TENURE≤2	Medium Tenure	*TENURE>7*	Total
NPTA < −2%	196 (43.9%)	178 (42.2%)	48 (31.6%)	422
NPTA −2 to 0% (just misses)	17 (3.8%)	23 (5.5%)	4 (2.6%)	44
NPTA 0 to 2% (just beats)	26 (5.8%)	22 (5.2%)	11 (7.2%)†	59
NPTA > 2%	208 (46.5%)	199 (47.1%)	89 (58.6%)	496
Total	447 (100%)	422 (100%)	152 (100%)	1,021

Panel B: Logistic Regression Results

$$\begin{aligned}\Pr(\mathit{BEATS\_BE} \text{ or } \mathit{MISSES\_BE} = 1) = {} & \beta_0 + \beta_1 \mathit{AUDFIRM} + \beta_2 \mathit{PBANK} + \beta_3 \mathit{SIZE} \\ & + \beta_4 \mathit{AGE} + \beta_5 \mathit{MKTVAL} + \beta_6 \mathit{MINING} \\ & + \beta_7 \mathit{TENURE}{\leq}2 + \beta_8 \mathit{TENURE}{>}7 + \varepsilon.\end{aligned}$$

	Just Beats Breakeven			Just Misses Breakeven		
	Exp. Sign	Coefficient	Sig.	Exp. Sign	Coefficient	Sig.
Experimental Variable						
TENURE > 7	+	.445	.134	−	−.778	.085*
Control Variables						
TENURE 2	+	−.147	.327	−	−.514	.073*
PBANK	+	−.047	.212	−	−.058	.086*
SIZE	−	.851	.000**	+	.225	.240
AUDFIRM	−	.224	.480	+	−.045	.893
AGE	−	−.355	.017**	+	−.137	.428
MKTVAL	+	−.730	.000**	−	−.256	.169
MINING	?	−.112	.744	?	.555	.102
Constant		−4.857	.001**		−2.416	.122
Adj. R^2 (Pseudo R^2)		.094			.032	

† Significant difference from medium tenure at the 5 percent level (one-tailed) for the proportion of just beats to just misses.

*, ** Significant at the 10 percent and 5 percent levels, respectively.

Variable Definitions:

NPTA = net profit to total assets;
BEATS_BE = 1 if profit is less than two percent of total assets, and 0 otherwise;
MISSES_BE = 1 if where loss is less than 2 percent of total assets, and 0 otherwise;
TENURE>7 = 1 if audit partner is engagement partner on a client company more than seven years, and 0 otherwise;
TENURE≤2 = 1 if audit partner is engagement partner on a client company less than or equal to two years, and 0 otherwise;
PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;
SIZE = the natural logarithm of total assets of the company at financial year-end;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
AGE = the natural logarithm of the number of years since the company was listed;
MKTVAL = the natural logarithm of the market value of equity; and
MINING = 1 if mining industry, and 0 otherwise.

TABLE 5
Just Beats and Just Misses Last Year's Profit
(n = 965)

Panel A: Earnings Benchmark Frequencies and Statistics

	TENURE≤2	Medium Tenure	*TENURE>7*	Total
$\Delta NPTA < -2\%$	142 (36.3%)	149 (35.3%)	45 (29.6%)	336
$\Delta NPTA$ −2 to 0% (just misses)	47 (12.0%)	49 (11.6%)	15 (9.9%)	111
$\Delta NPTA$ 0 to 2% (just beats)	59 (15.1%)†	94 (22.3%)	37 (24.3%)	190
$\Delta NPTA > 2\%$	143 (36.6%)	130 (30.8%)	55 (36.2%)	328
Total	391 (100%)	422 (100%)	152 (100%)	965

Panel B: Logistic Regression Results

$$\Pr(BEATS\_LYR \text{ or } MISSES\_LYR = 1) = \beta_0 + \beta_1 AUDFIRM + \beta_2 PBANK + \beta_3 SIZE + \beta_4 AGE + \beta_5 MKTVAL + \beta_6 MINING + \beta_7 TENURE{\leq}2 + \beta_8 TENURE{>}7 + \varepsilon.$$

	Just Beats Last Year			Just Misses Last Year		
	Exp. Sign	Coefficient	Sig.	Exp. Sign	Coefficient	Sig.
Experimental Variable						
TENURE > 7	+	−.111	.329	−	−.302	.175
Control Variables						
TENURE ≤ 2	+	−.337	.047**	−	.093	.342
PBANK	+	−.034	.165	−	−.034	.238
SIZE	−	.188	.106	+	.512	.000**
AUDFIRM	−	−.276	.172	+	−.023	.920
AGE	−	.096	.345	+	.041	.731
MKTVAL	+	.201	.078*	−	−.338	.006**
MINING	?	−1.003	.000**		.162	.496
Constant		−7.853	.000**		−5.415	.000**
Adj. R^2 (Pseudo R^2)		.211			.051	

---

† Significant difference from medium tenure at the 5 percent level (one-tailed) for the proportion of just beats to just misses.

*, ** Significant at the 10 percent and 5 percent levels, respectively.

Variable Definitions:

ΔNPTA = change from last year in net profit to total assets;
BEATS_LYR = 1 if increase over last year's earnings is less than 2 percent of total assets, and 0 otherwise;
MISSES_LYR = 1 if decrease from last year's earnings is less than 2 percent of total assets, and 0 otherwise;
TENURE>7 = 1 if audit partner is engagement partner on a client company more than seven years, and 0 otherwise;
TENURE≤2 = 1 if audit partner is engagement partner on a client company less than or equal to two years, and 0 otherwise;
PBANK = probability of bankruptcy as measured by adjusted Zmijewski score;
SIZE = the natural logarithm of total assets of the company at financial year-end;
AUDFIRM = 1 if audited by Big 6 firm, and 0 otherwise;
AGE = the natural logarithm of the number of years since the company was listed;
MKTVAL = the natural logarithm of the market value of equity; and
MINING = 1 if mining industry, and 0 otherwise.

last year's profit figure. For short partner tenure the proportion of the combined just beats and just misses observations that do just beat last year's profit is 59 of 106 (55.7 percent). For medium tenure this is 94 of 143 (65.7 percent), and for long tenure this is 37 of 52 (71.2 percent). The proportion of long tenure versus medium tenure observations is not significant, suggesting no diminution in audit quality associated with long audit partner tenure. Separate regression analyses for the just beats (just misses) last year's profit are contained in Table 5, Panel B. Results also do not support the hypothesized deterioration in audit quality associated with long audit partner tenure, with *TENURE>7* not significant for either the "beats last year" or "misses last years" profit.

Sensitivity Analyses

The beats/misses results provide limited support for the hypothesized deterioration in audit quality with long audit partner tenure. This may be caused by the small number of long tenure observations for the *TENURE>7* variable. Substituting *TENURE>7* with a binary variable measuring partner tenure greater than five years (*TENURE>5*), long tenure is no longer approaching significance in the *BEATS* or *MISSES_BE* analyses (and continues to be not significant in the *BEATS* or *MISSES_LYR* analyses). This result suggests that observations for partner tenure of six or seven years behave more like medium than long tenure observations.

To further investigate if the propensity for company earnings to just beat or miss specific benchmarks is associated with audit partner tenure, we substitute the two tenure variables (*TENURE≤2*, *TENURE>7*) with a single linear *TENURE* variable. We also control for instances of possible lower audit quality in the first two years of partner tenure by excluding the 447 *TENURE≤2* observations. For the just beats last year's profit analysis, we find the effect of *TENURE* is positive and approaching significance ($p = .077$, one-tailed) while it remains insignificant for the other three benchmark analyses.[20]

To explore the potential interaction between long partner tenure and *AUDFIRM*, we replicate the approach used for the going-concern audit opinion and *AWCA* analyses. Excluding first- and second-year tenure observations and running *TENURE* as a continuous variable in place of *TENURE>7*, we find a significant *AUDFIRM*TENURE* interaction for one of the analyses, the just beats last year's profit regression. *TENURE* is positive and significant ($p = .008$, one-tailed) and *AUDFIRM*TENURE* is negative and significant ($p = .010$, one-tailed). This suggests that non-Big 6 audit firm clients are more likely to just beat last year's profit.[21] The interaction term *AUDFIRM*TENURE* is insignificant for the other three analyses.

Further sensitivity analyses are undertaken where the cutoff for *BEATS* or *MISSES* breakeven and prior-year earnings is changed from 2 percent to either 1 or 3 percent. Varying the cutoff to either 1 or 3 percent produces similar results to the primary analyses.

In summary, the results of our analysis of the distribution of reported earnings provides some evidence suggesting a deterioration in audit quality associated with long audit partner tenure. Consistent with results from the going concern analysis (Section VI), sensitivity analysis around beats (misses) last year's profit suggests the deterioration in audit quality is associated with the non-Big 6 firms.

---

[20] When we square the tenure variable to accentuate the effect of partner tenure, the coefficients for *TENURE* do not change in significance.

[21] For the 94 of 143 medium tenure observations in the range of +/−2 percent Δ*NPTA*, 70 of 101 (69 percent) of the companies audited by the Big 6 firms just beat last year's *NPTA*, compared with 24 of 42 (57 percent) for companies audited by non-Big 6 firms. For the 37 of 52 long tenure observations, 16 of 26 (62 percent) of the companies audited by the Big 6 firms just beat last year's *NPTA*, compared with 21 of 26 (81 percent) for companies audited by non-Big 6 firms.

IX. OVERALL SENSITIVITY ANALYSIS: COMPARISON OF THE THREE METHODS USING CONSTANT SAMPLE

One of the possible reasons for the difference in the results between the three methods of analyses is the variations in sample size. We therefore rerun all the analyses using the sample of 965 used in the just beating or missing last year's earnings, which involves deleting the 56 new listings from our sample of 1,021. We find that under this approach *TENURE*>*7* remains significant for the going concern analysis (p = .038, one-tailed) for the reduced sample of 532 distressed companies contained in the sample of 965. With regard to the accruals models, while we are unable to rerun the *AWCA* because it requires last year's sales as a base level of activity, we are able to run the current accruals model on the 965 observations. We find that *TENURE*>*7* remains insignificant for both the signed and absolute measures of current accruals. With regard to the earnings benchmarks for breakeven, there is little difference in the result, with *TENURE*>*7* remaining negative for the just misses (p = .087, one-tailed) and positive but insignificant for the just beats. It therefore appears that the differences in findings for the different measures of audit quality are not a result of different sample compositions.

X. CONCLUSION

One of the main policy initiatives that has been emphasized in recent times in order to address concerns about audit quality is mandatory rotation of the audit partner. This study provides evidence on the underlying rationale for an audit partner rotation policy: there is a reduction in audit quality associated with long tenure. Examining partner tenure in an environment where the name of the engagement partner is publicly disclosed on the audit report and at a time when a policy of partner rotation is not mandated allows for an empirical examination of evidence to support this rationale. We examine 1995, a time before audit partner tenure became a professional, regulatory or audit firm requirement in Australia.

We examine the relationship between audit partner tenure and three proxies for audit quality. These are, first, auditors' propensity to issue going-concern audit opinions for distressed companies; second, an examination of the signed and absolute amounts of abnormal working capital accruals; and third, the extent to which key earnings benchmarks are just beaten and just missed.

In examining going-concern modified opinions, we find that the auditor's propensity to issue such opinions do diminish over the audit partner's tenure, after controlling for other factors which influence this propensity. Sensitivity analyses suggest that the decrease in propensity to qualify with long tenure is associated with non-Big 6 firms.

We find no evidence of an association of either the signed or absolute amount of abnormal working capital accruals with long audit partner tenure. This finding is inconsistent with Chi and Huang (2005) who find some support for the view that long audit partner tenure is associated with reduced earnings quality. Although both studies are undertaken in a pre-mandatory period, these differences may be a result of cultural disparities, and variations in the measurement of audit partner tenure and earnings quality.

In examining key earnings targets we find some evidence of just beating (missing) earnings benchmarks for long tenure observations. A lower proportion of clients miss breakeven for long partner tenure observations, suggesting a greater ability to manage earnings in order to report a profit during later years of partner tenure. Also, sensitivity analysis around beats (misses) last year's reported profit suggests deteriorating audit quality is more likely to be associated with the clients of non-Big 6 firms. When these results are combined with the findings of the abnormal working capital accruals analysis, it appears that mechanisms other than working capital accruals variations are being used to achieve these outcomes.

In conclusion, for the measures of audit quality examining the auditor's propensity to issue a going-concern audit opinion and just meeting (missing) earnings benchmarks, we find evidence consistent with a diminution in audit quality associated with long audit partner tenure. These results provide support for the introduction of a rotation policy after seven years of partner tenure. Sensitivity analyses show that the evidence of deterioration in audit quality seems to pertain mainly to non-Big 6 auditors. This is consistent with prior studies that provide evidence of higher quality audits provided by Big 6 firms (for example, Francis et al. 1999; Becker et al. 1998).

The analysis is conducted during a period before there was mandatory rotation of audit partners. Therefore, we cannot necessarily expect that outcomes of the audit process observed over a long period of tenure are the same as those observed under any legislated period (i.e., five or seven years) of limited tenure. Our study is also not relevant to the consideration of whether a shorter five-year rotation period should be adopted, as has recently occurred in many jurisdictions. It is possible that the behavior of audit partners might be impacted by the knowledge that their tenure as audit partner will end at a specified date. Future research might investigate the impact on audit quality of the introduction of such a policy and any variations in its requirements.

REFERENCES

American Institute of Certified Public Accountants (AICPA), Commission on Auditors' Responsibilities (Cohen Commission). 1978. *Commission on Auditors' Responsibilities: Report, Conclusions and Recommendations.* New York, NY: AICPA.

———, SEC Practice Section. 1992. *Statement of Position Regarding Mandatory Rotation of Audit Firms of Publicly Held Companies.* New York, NY: AICPA.

Arrunada, B., and C. Paz-Ares. 1997. Mandatory rotation of company auditors: A critical examination. *International Review of Law and Economics* 17 (March): 31–61.

Ashbaugh, H., R. LaFond, and B. W. Mayhew. 2003. Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review* 78 (July): 611–640.

Australian Commonwealth Government, Audit Review Working Party. 1996. *Review of the Requirements for the Registration and Regulation of Auditors.* Canberra, Australia: Australian Commonwealth Government.

Australian Society of CPAs and The Institute of Chartered Accountants in Australia. 1994. *A Research Study on Financial Reporting and Auditing—Bridging the Expectation Gap.* Melbourne, Australia: Australian Society of CPAs.

Barth, M. E., J. A. Elliot, and M. W. Finn. 1999. Market rewards associated with patterns of increasing earnings. *Journal of Accounting Research* 37 (Autumn): 387–413.

Bazerman, M. H., K. P. Morgan, and G. F. Loewenstein. 1997. The impossibility of auditor independence. *Sloan Management Review* 38 (Summer): 89–94.

Becker, C. L., M. L. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (Spring): 1–24.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24 (December): 99–126.

Butterworth, S., and K. Houghton. 1995. Auditor switching and the pricing of audit services. *Journal of Business Finance and Accounting* 22 (April): 431–454.

Carcello, J. V., R. H. Hermanson, and H. F. Huss. 1995. Temporal changes in bankruptcy-related reporting. *Auditing: A Journal of Practice & Theory* 14 (Fall): 133–143.

Chi, W., and H. Huang. 2005. Discretionary accruals, audit-firm tenure and audit-partner tenure: Empirical evidence from Taiwan. *Journal of Contemporary Accounting and Economics* 1 (June): 65–92.

Committee on the Financial Aspects of Corporate Governance (Cadbury Committee). 1992. *Report of the Committee on the Financial Aspects of Corporate Governance.* London, U.K.: Gee and Co. Ltd.

Craswell, A. 1996. *Who Audits Australia?* University of Sydney, Australia: Accounting Foundation.

Davis, L. R., B. Soo, and G. Trompeter. 2003. Auditor tenure, auditor independence and earnings management. Working paper, Boston College, Boston, MA.

DeAngelo, L. E. 1981. Auditor independence, "low balling" and disclosure regulation. *Journal of Accounting and Economics* 3 (August): 113–127.

———. 1986. Accounting numbers as market valuation substitutes: A study of management buyouts of public stockholders. *The Accounting Review* 61 (July): 400–420.

Dechow, P. M., R. G. Sloan, and A. P. Sweeny. 1995. Detecting earnings management. *The Accounting Review* 70 (April): 193–226.

DeFond, M., and J. Jiambalvo. 1993. Factors related to auditor-client disagreements over income-increasing accounting methods. *Contemporary Accounting Research* 9 (Spring): 411–431.

———, and C. W. Park. 2001. The reversal of abnormal accruals and the market valuation of earnings surprises. *The Accounting Review* 76 (July): 375–404.

———, K. Raghunandan, and K. R. Subramanyam. 2002. Do nonaudit service fees impair auditor independence? Evidence from going-concern audit opinions. *Journal of Accounting Research* 40 (September): 1247–1274.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72 (January): 1–33.

Dopuch, N., R. R. King, and R. Schwartz. 2003. Independence in appearance and in fact: An experimental investigation. *Contemporary Accounting Research* 20 (Spring): 79–114.

Financial Reporting Commission (Ryan Commission). 1992. *Report of the Commission of Inquiry into the Expectations of Users of Published Financial Statements.* Dublin, Ireland: The Institute of Chartered Accountants in Ireland.

Francis, J. R., E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (Fall) 17–34.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement): 71–105.

Ghosh, A., and D. Moon. 2005. Does audit tenure impair audit quality? *The Accounting Review* 80 (April): 585–612.

Geiger, M., and K. Raghunandan. 2002. Auditor tenure and audit reporting failures. *Auditing: A Journal of Practice & Theory* 21 (March): 67–78.

Grant, J., R. Bricker, and R. Shiptsova. 1996. Audit quality and professional self-regulation: A social dilemma perspective and laboratory investigation. *Auditing: A Journal of Practice & Theory* 15 (Spring): 142–156.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (September): 125–53.

Holland, D., and A. Ramsay. 2003. Do Australian companies manage earnings to meet simple earnings benchmarks? *Accounting and Finance* 43 (March): 41–62.

Hopwood, W., J. McKeown, and J. Mutchler. 1989. A test of the incremental explanatory power of opinions qualified for consistency and uncertainty. *The Accounting Review* 64 (January): 28–48.

Hoyle, J. 1978. Mandatory auditor rotation: The arguments and an alternative. *The Journal of Accountancy* 145 (May): 69–78.

International Federation of Accountants (IFAC). 2003. *Code of Ethics*. New York, NY: IFAC.

Johnson, V. E., I. K. Khurana, and J. K. Reynolds. 2002. Audit firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19: 637–660.

Matsumoto, D. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77 (July): 483–514.

Mautz, R. K., and H. A. Sharaf. 1961. *The Philosophy of Auditing*. Sarasota, FL: American Accounting Association.

McLaren, N. L. 1958. Rotation of auditors. *The Journal of Accountancy* (July): 41–44.

Menon, K., and D. D. Williams. 2004. Former audit partners and abnormal accruals. *The Accounting Review* 79 (October): 1095–1118.

Mutchler, J. F., W. Hopwood, and J. C. McKeown. 1997. The influence of contrary information and mitigating factors on audit opinion decisions on bankrupt firms. *Journal of Accounting Research* 35 (Autumn): 295–310.

Myers, J., L. A. Myers, and T. C. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (July): 779–800.

Nelson, M. W., J. A. Elliot, and R. L. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings-management decisions. *The Accounting Review* (Supplement): 175–202.

Reynolds, J. K., and J. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30: 375–400.

Shockley, R. A. 1981. Perceptions of auditors' independence: an empirical analysis. *The Accounting Review* 56 (October): 785–800.

St. Pierre, K., and J. A. Anderson. 1984. An analysis of factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 243–263.

U.S. Government Accounting Office (GAO). 2003. *Public Accounting Firms: Required Study on the Potential Effects of Mandatory Audit Firm Rotation*. Report to the Senate Committee on Banking, Housing, and Urban Affairs and the House Committee on Financial Services. November. Washington, D.C.: Government Printing Office.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 50–82.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 677–712

Optimal Information Asymmetry

Madhav V. Rajan
Stanford University

Richard E. Saouma
Stanford University

ABSTRACT: At the heart of decentralization lies the notion that tasks are delegated by owners to managers who possess superior local information. The extent of this information asymmetry is often an endogenous construct, as it is influenced by the owner's choice of internal accounting systems and the manager's investment in acquiring local expertise. In this paper, we explore how varying levels of pre-contract, asymmetric information affect the owner-manager relationship. We provide three main sets of insights. First, we find that the owner's payoffs are initially decreasing, and strictly convex everywhere, in the quality of the manager's private information. The owner thus prefers to deal with either a perfectly informed or a perfectly uninformed manager, and we characterize conditions for either to be the preferred choice. Second, in contrast to recent work, we demonstrate that when information can be communicated internally, the optimal strength of managerial incentives unambiguously decreases as the manager becomes better informed. Third, we derive the surprising result that a self-interested manager does not always prefer to maximize his informational advantage. Our work has implications for the optimal design of organizations, and for internal accounting and control systems in particular.

I. INTRODUCTION

An issue of central interest in management accounting is the choice of information systems to aid managers in their decision making. In decentralized firms, decisions are delegated to managers who may have information superior to that of the owners of the firm. Some of this information superiority arises due to the manager's expertise regarding local conditions, such as his knowledge of consumer preferences in specific markets or for specific product attributes. However, the relative information advantage of the manager is also a function of investments made by the firm in internal information systems (e.g., the installation of ERP systems or the acquisition of market survey data), as well as investments made by the manager in acquiring greater local expertise. In settings where the information generated by these systems can be monitored by the owner, it is clear that, holding cost considerations aside, more information is better. However, in decentralized

---

We thank Stan Baiman, Tim Baldenius, Paul Fischer, Alexander Nezlobin, Stefan Reichelstein, John Roberts, Bob Wilson, and an anonymous reviewer for helpful comments. We are indebted to Raffi Indjejikian (editor), to participants in the informal Accounting and Economics workshops at Stanford University, and to Harry Evans, our discussant at the 2005 AAA meetings, for many constructive suggestions. Professor Saouma gratefully acknowledges funding from the KPMG Foundation.

Editor's note: This paper was accepted by Raffi Indjejikian.

Submitted March 2005
Accepted November 2005

environments where the realized information signals are privately observed by managers, owners may prefer to deal with managers whose information is less than perfect. The reason is that while managers can potentially use additional information to make better decisions for the firm, inducing them to do so against their self-interest may require excessive levels of additional compensation.

In this paper, we analyze the effects of contracting with decentralized managers who possess varying levels of pre-contract informational superiority. We first propose a measure of informational quality and validate it by demonstrating that first-best profit levels are increasing and strictly convex in this metric. We then analyze the efficient second-best outcome and show that the owner's profit is initially decreasing and strictly convex everywhere as a function of the manager's private information quality. An immediate implication is that the firm's owner always prefers for the manager to be either perfectly informed or perfectly uninformed. Ignoring cost considerations, the owner will thus never offer the manager some, but not all of the information available. Moreover, giving an uninformed manger an infinitesimal amount of additional information is always detrimental to the owner. Whether the owner will prefer for the manager to be perfectly informed or perfectly uninformed depends on how much the underlying states of the world vary. When the states of the world vary greatly or there is a lot of uncertainty regarding the actual state, we show that the benefits to informing the manager are sufficient to overcome any accompanying informational rents. On the other hand, when the states of nature are sufficiently similar, the owner chooses not to inform the manager because this precludes the latter from earning any informational rents, whereas the consequences of the manager making an uninformed decision are negligible.

We show that the above results hold regardless of whether the manager's contractible output is deterministic or stochastic. In both settings, we allow for full communication of the manager's information using menus of contracts. In the stochastic model, we employ menus of linear contracts, which allows us to measure the effect of information on incentive strength. In contrast to recent work that has studied this issue in settings with no communication, we show that the optimal strength of managerial incentives always decreases as the manager's level of private information increases.

We derive the rather unexpected finding that the manager's informational rents do not necessarily increase as a function of his informational advantage. Both in the deterministic and stochastic settings, once the manager's level of private information is sufficiently large, his incentive to lie to the owner decreases, and the owner can reduce the level of rents used to keep the manager honest. We further show that while the owner and the manager would never select the same level of private information quality *ex ante*, the manager's desired choice at times more closely approximates the welfare-maximizing level of private information quality. We thus identify the existence of settings in which there is an endogenous demand for the owner to delegate or sell the right to improve informational quality to the manager.

Given the importance of information system choice in managerial accounting, it is not surprising that there is a sizable extant literature on this topic. The vast majority of papers, however, have looked at the choice of noisy performance measures to *ex post* evaluate the hidden actions of risk-averse managers (Banker and Datar 1989; Datar et al. 2001). In this setting, if the principal can commit to his use of information, then the informativeness principle (Holmstrom 1979) generally allows one to determine whether additional *ex post* information is of value.[1] Unfortunately, there is no analogous condition to value *ex ante* information. The accounting literature first examined the value of different information systems within an agency construct in Christensen (1981). Via a series of examples, Christensen (1981) finds that offering the agent more post-contract, pre-decision information (in the sense of Blackwell [1951]) may not necessarily make the principal better off. Following Christensen's (1981) seminal contribution that more pre-decision information may be detrimental, the majority of the follow-on accounting literature has searched for sufficient conditions for making such information beneficial (see, e.g., Baiman and Evans 1983; Penno 1984; Baiman and Sivaramakrishnan 1991). In all of these models, the agent is risk-averse, and hence the principal tempers efficiency with the cost of paying the agent informational rents and risk premia. Because we are primarily interested in informational preferences, we assume all parties are risk-neutral, and examine the principal and agent's preferred levels of informational advantage. With risk-neutrality, if the agent's information arrives after signing the contract, then the principal can obtain the first-best outcome by "selling" the agent the firm. To rule out such pathological cases, we assume that the agent faces limited liability and possesses pre-contract information, as in Sappington (1984).

The economics literature has studied pre-decision information with risk-neutral agents (e.g., Lewis and Sappington 1991, 1993; Cremer and Khalil 1992; Cremer et al. 1998; Kessler 1998). This stream of work generally assumes that the onus to become asymmetrically informed rests on the agent; the bulk of the analysis hence rests in providing incentives for the agent to seek pre-decision information and examining the ensuing equilibria. Unlike the economics literature, we allow for the possibility that the principal can influence the agent's level of private information. Two other important distinctions are that we allow the agent to be imperfectly informed (which we show in many cases is his preferred choice), and we specify the quality of the agent's private information to be common knowledge.

More closely related to our paper are recent studies that have looked at the effect of private information on incentives (Bushman et al. 2000; Baker and Jorgensen 2003). Both model the presence of imperfectly informed agents, but neither allows for contracting based on reported information. In particular, Baker and Jorgensen (2003) show that as an agent becomes more informed, the owner will want to raise the agent's incentives. We show that in a full contracting framework with communication, the owner's optimal response is rather to *decrease* the strength of incentives as the manager becomes more informed. In Bushman et al. (2000), the authors compare the principal's payoffs when the agent is informed and his actions are unobservable (delegation) to the case where the agent is uninformed and his actions are observable (centralization). The authors show that although congruence between the agent's performance measure and his individual task is beneficial, excessive congruence makes delegation less appealing. Both Bushman et al. (2000) and Baker and Jorgenson (2003) assume that the agent's information is post-contract, that is, the agent's individual rationality constraint must be met in expectation over all states of the world. We assume that the agent's information is pre-contract, such that the contract offered by the principal must satisfy the agent's individual rationality constraint state-by-state. A recent paper examining the effect of imperfect, pre-contract information is Raith (2004), which asks whether compensation should be based on noisy output when the agent's effort is perfectly contractible. Raith (2004) finds that as the agent's signal becomes more informative, the owner will place more emphasis on noisy output and less emphasis on effort. We do not assume that the agent's effort is contractible or even verifiable. Moreover, while in Raith's (2004) two-state model, the owner's expected profits are increasing in the agent's pre-decision information, we find that the owner's expected profits are initially decreasing in the agent's level of private information and convex overall.

---

[1] For a nice illustration of the failure of the informativeness principle when the principal cannot commit, see Arya et al. (1997).

In its examination of the owner and agent's divergent preferences over private information, our study is most similar to Antle and Fellingham (1995). Using a model of resource allocation with hurdle rates, the authors seek to determine the extent to which the owner and manager would benefit from making the agent's private information public. Their resource allocation model has a single hurdle rate, but several states of nature. Thus, communication between the owner and agent is coarse[2] and allows the agent to obtain positive rents in good states of the world. The authors find that divulging some of the agent's private information can either increase or decrease the agent's informational rent.[3] In their model, the agent's rents increase whenever the owner operates in states that were deemed unprofitable without the agent's private information, allowing the agent to capture rents in more states. In our model, we vary the agent's private information by altering the correlation between the agent's private signal and the state of the world. As such, no new contracting opportunities arise as the agent's private information varies.

The paper proceeds as follows. In Section II, we introduce and fully analyze the owner's preferences over information when the manager's output is deterministic. In Section III, we generalize our findings to the case of noisy output, using menus of linear contracts. Section IV examines the agent's informational preferences and the welfare-maximizing level of private information quality. Section V concludes.

II. MODEL

We model an agency relationship where all actions have been delegated by the shareholders (principal/owner) to a manager (agent). We assume that the manager has distinct tasks to perform. If he chooses effort level e_j on task j, then his total disutility is given by $c(\mathbf{e})$ where $\mathbf{e} = \{e_1,\ldots,e_k\}$. The manager's effort translates into revenues π for the owner according to the formula:

$$\pi = \sum_{j=1}^{k} \beta_j e_j$$

where the vector $\boldsymbol{\beta} = \{\beta_1,\ldots,\beta_k\}$ is common knowledge. To simplify our analysis, we assume that $c(\mathbf{e}) = \frac{\sigma}{2}\sum_{j=1}^{k} e_j^2$, where the parameter σ is a random variable with support $\{\sigma_1,\ldots,\sigma_n\}$. Since σ measures the cost of effort, larger values of σ correspond to less productive managers, whereas smaller of values of σ denotes good "types" of managers.[4] At the time of contracting, the value of σ is unknown to both the owner and manager. The owner and manager know only that σ has support $\{\sigma_1,\ldots,\sigma_n\}$ with corresponding probabilities $\{f(\sigma_1),\ldots,f(\sigma_n)\}$. However, before signing a contract and selecting his effort, we assume that the manager has private access to a signal $s \in \{s_1,\ldots,s_n\}$, which is related to σ as follows:

---

[2] The communication is in fact binary—whether the agent's productivity is above or below the hurdle rate.

[3] See Antle and Fellingham (1995, Examples 1 and 2).

[4] All of our results continue to hold if we assume that the agent's private information concerns the agent's output (revenue) rather than his costs.

$$\Pr[\sigma = \sigma_j | s = s_i] = \begin{cases} a + (1 - a)f(\sigma_j) & i = j \\ (1 - a)f(\sigma_j) & i \neq j \end{cases}$$

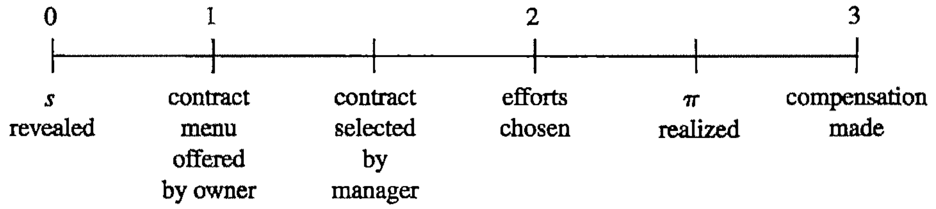
where $a \in [0,1]$ is an informational parameter. The conditional density of σ given s implies that the unconditional distribution of s is identical to σ, that is $f(\sigma_i) = f(s_i)$ for $i = 1,\ldots,n$. The following Lemma relates a to s and σ.

Lemma 1: The correlation between s and σ is given by a.

Clearly, the manager's information is increasing in a, hence we shall say that the manager is more informed for larger values of a and less informed for smaller values of a. Note that if $a = 0$, then the signal s is uncorrelated with the true state of the world, σ, and is thus entirely without information content. When $a = 0$, we shall say that the manager is perfectly uninformed. On the other hand, if $a = 1$, then s perfectly reveals σ, in which case we label the manager as perfectly informed. The case of $a = 1$ coincides with the adverse-selection paradigm originating in Maskin and Riley (1984). We will initially assume that a is exogenously specified, but will later treat a as a choice variable.[5]

The sequence of events is as follows. At time 0, the manager receives s, and at time 1 the owner offers the manager a menu of contracts. If the manager accepts, then the manager communicates his choice of contracts from the menu to the owner. If the manager declines, then we assume both parties earn 0, and the relationship ends. At time 2, the manager selects his effort levels and the owner consequently earns π. At time 3, the owner pays the manager according to the terms of the initial contract. Figure 1 illustrates the entire sequence of events.

FIGURE 1
Time Line



---

[5] To illustrate our informational assumptions, consider a two-state example where the unconditional density is uniform. Then, the conditional probabilities in our specification reduce to the following intuitive expressions:

	$\Pr[\sigma_1 \| s]$	$\Pr[\sigma_2 \| s]$
s_1	$\frac{(1 + a)}{2}$	$\frac{(1 - a)}{2}$
s_2	$\frac{(1 - a)}{2}$	$\frac{(1 + a)}{2}$

The fact that the manager receives s prior to the contract being signed (pre-contract information) is an important feature of our model. As a natural counterpart to this specification, we stipulate that any contract must guarantee the manager at least a minimum level of utility (which we set at zero without loss of generality) for every possible realization of s. Were the contract to be signed before the manager receives s (post-contract information), it is possible that the owner could "sell" the firm to the manager for an amount equal to the expected profits of the firm assuming the manager realizes the mean cost $\overline{\sigma}$. For every realization of s, the manager would then select the conditional first-best action, and there would be no scope for communication. However, note that in this setting, the manager's limited liability constraint would be violated for high (bad) signal realizations. On the other hand, if the owner cannot "sell" the firm to the manager *ex ante*, then all of our results transfer over to a post-contract information setting, because the owner faces the same constraints and objectives as in our model.

We first consider the benchmark where the signal s is assumed to be jointly observable and contractible. To simplify notation, let $\hat{\sigma}_i = \mathrm{E}[\sigma | s = s_i]$ and let $e_{i,j}$ denote the effort the manager will put forth on task j when the public signal s_i is observed. Given our conditional probability structure, we can write $\hat{\sigma}_i(a) = a\sigma_i + (1 - a)\overline{\sigma}$. In this setting, when the owner observes signal s_i and pays the manager p_i conditional on observing the correct level of revenues, she solves:

$$\max_{e_{i,j}p_i} \sum_{j=1}^{k} e_{i,j}\beta_i - p_i$$

such that:

$$p_i - \frac{\hat{\sigma}_i}{2}\sum_{j=1}^{k} e_{i,j}^2 \geq 0.$$

The owner will clearly leave the manager with his reservation utility of zero in every state by setting $p_i = \frac{\hat{\sigma}_i}{2}\sum_{j=1}^{k} e_{i,j}^2$. This reduces the owner's problem to:

$$\max_{e_{i,j}} \sum_{j=1}^{k} e_{i,j}\beta_i - \frac{\hat{\sigma}_i}{2}\sum_{j=1}^{k} e_{i,j}^2.$$

Solving this, the owner sets:

$$e_{i,j}^{fb} = \frac{\beta_j}{\hat{\sigma}_i}$$

and we can thus write the owner's expected profits as:

$$\pi^{fb}(a) = \frac{(\beta \cdot \beta)}{2}\sum_{i=1}^{n} \frac{f(\sigma_i)}{\hat{\sigma}_i(a)}.$$

Note that the expected profits vary over a only via the (conditional) expected states of the world $\hat{\sigma}_i(a)$. When the signals are uninformative, the conditional expected value of σ is simply the unconditional mean $\bar{\sigma}$, that is $\hat{\sigma}_i(0) = \bar{\sigma}$ for $i = 1,\ldots,n$. When the signal is perfectly informative, the signal perfectly reveals the state of the world, that is $\hat{\sigma}_i(1) = \sigma_i$. The following proposition demonstrates that in this setting, better information translates unambiguously to increased profits.

Proposition 1: First-best profits are increasing and strictly convex in a.

When a increases, the amount of information contained in the signal increases. In the first-best setting, the owner costlessly observes the signal s and can then make better decisions as to how much effort to ask of the manager. Higher levels of a thus result in higher levels of profits, as well as higher marginal profits. Note that the above result also serves to validate our choice of a as a parametrization of the level of information available.

In a second-best setting, where only the manager learns s, it is less clear that the owner would always prefer higher levels of a. In particular, the owner may have to pay the manager informational rents in order for the manager to truthfully reveal (or make optimal use of) his private signal s.[6]

We now turn to the owner's problem when only the manager can observe the signal s, and the manager's output is deterministic. When revenues are deterministic, a mechanism where the manager reveals his information to the owner and the owner offers a contract on the basis of the revelation is equivalent to a mechanism without communication whereby the owner offers the manager a menu of revenue targets and corresponding payments. We assume that the owner employs the latter type of contract, and in particular specifies n {*revenue, payment*} pairs, $\{r_i, p_i\}$, from which the manager must choose.

When the manager privately observes signal $s = s_i$ and "targets" revenue level r, he chooses his effort allocation across tasks as the solution to the following problem:

$$\min_{\mathbf{e}_i} \frac{\hat{\sigma}_i}{2} \sum_{j=1}^{k} e_{i,j}^2$$

such that:

$$\boldsymbol{\beta} \cdot \mathbf{e}_i = r,$$

where $\mathbf{e}_i = \{e_{i,1},\ldots,e_{i,n}\}$ is the effort vector chosen by a manager with signal $s = s_i$. Solving the above program, we find that the manager will expend effort on task j of:

$$e_{i,j} = \frac{\beta_j}{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})}\, r.$$

Let $c_i(r)$ denote the minimum (expected) cost for a type i manager to generate r profits. Given the optimal effort allocation, we find that:

---

[6] In a post-contract information setting where the manager is committed to stay with the firm and has no minimum utility requirement, the owner would not have to pay any informational rents. In such a setting, the result in Proposition 1 implies that the owner would always prefer, if given the choice, to have as high a level of a as possible.

$$c_i(r) = \frac{\hat{\sigma}_i}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} r^2.$$

Given this expected cost of attaining a revenue target as a function of the manager's signal, we can now specify the owner's optimization problem:

$$\max_{\mathbf{r},\mathbf{p}} \sum_{i=1}^{n} f(\sigma_i)(r_i - p_i) \tag{P1}$$

such that:

$$p_i - c_i(r_i) \geq 0 \quad \forall i \tag{1}$$

$$p_i - c_i(r_i) \geq p_l - c_i(r_l) \quad \forall i,l. \tag{2}$$

Constraint (1) ensures that the manager will accept the contract offered at time 1, regardless of the signal he obtained at time 0. Constraint (2) makes sure that the manager will truthfully reveal his signal to the owner by his choice of the appropriate {*revenue, payment*} pair.

To give a meaningful answer to the owner's problem, we make two additional assumptions:

A1: The probabilities $f(\cdot)$ satisfy the following condition:

$$\frac{F(\sigma_i)}{f(\sigma_i)} \text{ is increasing in } i,$$

where $F(\sigma_i)$ represents the cumulative distribution function.

A2: The states σ_i exhibit non-decreasing differences, i.e., $(\sigma_i - \sigma_{i-1})$ is weakly increasing in i.

Assumption 1 is standard and is satisfied by most common distributions, including the uniform distribution. It is the discrete analog to the usual assumption of log-concave distribution functions (see Bagnoli and Bergstrom 2005). Assumption 2 asks that the relative dispersion of the "bad" states be at least as great as that of the "good" states. Because an agent with signal s_i is tempted to report signal s_{i+1}, A2 bounds each agent's temptation to misreport by his neighbor's temptation. Note that A1 places a restriction on the probability weights, whereas A2 restricts the support of σ. Assumption 1 is used to prevent any pooling among the different types of managers, whereas A2 ensures that the managers with better signals are asked to produce more than those with less productive signals. A natural candidate for σ satisfying A1 and A2 is when the states of the world are equally spaced and uniformly distributed.

We can now solve (P1). If the *ex ante* quality of the information signal is a, then let $r_i(a)$ denote the optimal revenue targets issued to the manager when he reports signal $s = s_i$.

Proposition 2: Assuming A1 and A2, the solution to (P1) is given by:

$$r_i(a) = \frac{(\beta \cdot \beta)\hat{\sigma}_{i-1}^{-1}(a)}{\frac{F(\sigma_i)}{f(\sigma_i)}\left(\frac{\hat{\sigma}_i(a)}{\hat{\sigma}_{i-1}(a)} - 1\right) + 1};$$

$$p_i(a) = p_n(a) + \sum_{l=i}^{n-1} \frac{\hat{\sigma}_l(a)}{2(\beta \cdot \beta)} \left(r_l^2(a) - r_{l+1}^2(a)\right);$$

$$p_n(a) = c_n(r_n(a)).$$

The solution to the second-best problem depends on a exclusively via $\hat{\sigma}(a)$. Recall that the expected state of the world conditional on $s = s_i$, $\hat{\sigma}_i(a)$, is linear in a with endpoints $\hat{\sigma}_i(0) = \bar{\sigma}$ and $\hat{\sigma}_i(1) = \sigma_i$. When $a = 0$, all the revenue targets are the same and the owner (in effect) offers a single contract for all different types. As a increases, the n different contracts grow further apart. Following the monotonicity of types, the target revenues are decreasing in i, such that good types are asked to generate more revenue than worse types. The worst type ($s = s_n$) earns no rents, and the best type ($s = s_1$) puts forth the same level of effort as in the first-best setting and earns positive rents. Moreover, the informational rents paid to the manager are increasing in type. The owner's profits are increasing in the manager's type, but also vary over the level of private information contained in the manager's signal. The following proposition characterizes the effect of changes in the manager's private information on the owner's expected profits:

Proposition 3: Second-best profits are decreasing in a at $a = 0$, and are strictly convex over the entire range of a.

When $a = 0$, the owner offers a single contract and the manager earns no informational rents. As a increases and the contracts in the optimal menu grow further apart (the revenue targets diverge as do the payments), the manager begins to earn rents. Although increasing the manager's information helps him implement better decisions, for values of a sufficiently close to zero, these benefits are insufficient to warrant the higher rents necessary for the manager to truthfully reveal his signal to the owner. However, as a increases, the marginal benefit of informing the manager trumps the owner's marginal cost associated with paying informational rents; hence, the owner's profits are convex in a. Recall that in the first-best setting of Proposition 1, the value of additional information was increasing in the pre-existing level of information. Although it is impossible to disentangle the agency costs and the efficiency gains resulting from a change in a, Proposition 3 shows that the agency cost is sufficiently small that convexity in the principal's profits is preserved in the second-best setting. Moreover, Proposition 1 showed that the marginal benefit to the principal of increasing a when $a = 0$ is 0. Because the first- and second-best contracts are identical when $a = 0$, we can conclude that even when the agent is perfectly uninformed, there exists a positive agency cost associated with any increase in a. The agency costs associated with an informed agent and positive a are two-fold: the principal is paying the agent informational rents, but additionally, she distorts the agent's effort levels away from first-best in order to minimize the informational rents.[7]

An immediate implication of convexity is that, ignoring the costs of information acquisition, the owner will want the manager to either be perfectly informed ($a = 1$) or perfectly uninformed ($a = 0$).[8] An obvious question that arises is whether we can characterize settings in which the owner prefers an informed to an uninformed manager. The following proposition identifies a necessary and sufficient condition for the owner to prefer an informed manager to an uninformed one.

---

[7] In Section IV, we further examine the effect of varying a on the agent's informational rents.

Proposition 4: If the owner can choose the level of information conveyed in s, she will make s perfectly informative if and only if:

$$\frac{1}{\bar{\sigma}} \leq \mathrm{E}\left[\frac{\hat{\sigma}_{i-1}^{-1}}{\frac{F(\sigma_i)}{f(\sigma_i)}\left(\frac{\sigma_i}{\sigma_{i-1}} - 1\right) + 1}\right] \tag{3}$$

or, equivalently:

$$\frac{\beta \cdot \beta}{\bar{\sigma}} \leq \mathrm{E}[r(1)] \tag{4}$$

is satisfied.

To understand Condition (4), it is instructive to identify circumstances when it will not be met. When the agent's signal is perfectly uninformative, the principal earns $(\beta \cdot \beta)\frac{1}{2\bar{\sigma}}$. On the other hand, when the principal decides to offer the agent a perfectly informative signal, her expected profit is given by half of the agent's expected revenue target. When the states of nature do not vary significantly, the informed agent's revenue target does not vary significantly from the uninformed agent's target of $(\beta \cdot \beta)\frac{1}{2\bar{\sigma}}$; however, the informed agent's targets have a distortion built in to minimize the agent's rents. As such, when the states of nature are relatively similar, $\frac{1}{2}\mathrm{E}[r(1)]$ will tend to be less than $(\beta \cdot \beta)\frac{1}{2\bar{\sigma}}$, because the incremental profits resulting from knowing s are insufficient to overcome the efficiency loss induced by the optimal contract.

The following corollary demonstrates the above intuition explicitly by considering a setting where the states of nature are uniformly distributed and grow according to a geometric series. In this case, we show that the owner will prefer an informed manager when and only when the states of the world are sufficiently far apart.

Corollary 1: Suppose that the states of the world are uniformly distributed with support $\{\sigma_1,\ldots,\sigma_n\}$ where $\sigma_i = \sigma_1\rho^{i-1}$ for $i \geq 1$. Then, the owner prefers the manager to be perfectly uninformed if $\rho < \hat{\rho}(n)$, and she prefers the manager to be perfectly informed if $\rho > \tilde{\rho}(n)$ for some $1 \leq \hat{\rho}(n) \leq \tilde{\rho}(n) \leq 2$.

---

[8] Lewis and Sappington (1991) derive a similar "all-or-nothing" result in a Baron and Myerson (1982) model where the principal controls the probability with which an agent receives a perfectly revealing signal, as opposed to a null signal. In their setup, imperfect information systems of the kind we study are not considered and, moreover, the principal relies upon an additional communication stage to determine whether the agent is perfectly informed.

When ρ is sufficiently close to 1 (recall that the states are increasing, hence $\rho \geq 1$), the states are relatively close to one another, and the owner prefers the manager to be uninformed. When ρ approaches 2, the states begin to differ greatly and the owner prefers the manager to be perfectly informed. Again, when the manager is uninformed, the owner offers a single generic contract that specifies a constant revenue target irrespective of the manager's signal. This works well when the states are relatively close to one another, as the generic target is never too distant from the optimal target the owner would have chosen had she known s. Moreover, when the owner offers a single contract, she does not need the manager to communicate s and thus avoids paying any informational rents. On the other hand, when the states vary significantly, a single contract would specify a revenue target that is "significantly non-optimal" for many realizations of s. As such, when the states differ greatly, the owner prefers to issue individual targets according to the manager's report, even when it involves paying the manager informational rents.

A similar result holds when the states are uniformly distributed and the supports grow according to an arithmetic series. In fact, the intuition behind the similarity of states and the owner's informational preferences can be generalized to any distribution of states. To do so formally, we must specify a metric that captures the extent to which the states are similar to one another. To this end, we next introduce the notion of linear mean-preserving contractions.

> **Linear Mean-Preserving Contraction:** For any random variable σ with mean $\overline{\sigma}$, support $\{\sigma_1,\ldots,\sigma_n\}$, and respective probabilities $(f(\sigma_1),\ldots,f(\sigma_n))$, a linear mean-preserving contraction (lmpc) is defined by a random variable σ_λ with support $\{\sigma_{\lambda_1},\ldots,\sigma_{\lambda_n}\}$ and respective probabilities $(f(\sigma_1),\ldots,f(\sigma_n))$, where $\sigma_{\lambda_i} = \lambda\sigma_i + (1 - \lambda)\overline{\sigma}$ for $\lambda \in [0,1]$.[9]

A linear mean-preserving contraction (lmpc) is the reverse of a mean-preserving spread, although we restrict the contraction to occur linearly. In particular, note that σ_λ has support $\overline{\sigma}$ when $\lambda = 0$ and σ_λ has the same support as σ when $\lambda = 1$. Let L_i denote the line with endpoints $\overline{\sigma}$ and σ_i. Then the i^{th} element in the support of σ_λ is a point on L_i. Moreover, this point moves closer to $\overline{\sigma}$ for small values of λ, and moves closer to σ_i for larger values of λ. We have plotted the supports of σ_λ as a function of λ in Figure 2. By considering linear mean-preserving contractions on an arbitrary random variable σ, we are in effect parametrizing the similarity between states, such that the random variable σ_λ is identical to σ if $\lambda = 1$, and as λ tends to 0, the support of σ_λ collapses to $\overline{\sigma}$. Because σ_λ collapses the support of σ linearly to the mean, the expected value of σ_λ is equal to $\overline{\sigma}$ for any value of λ. Moreover, because the expected state of the world is the only statistic that affects the owner's payoffs when the manager is perfectly uninformed, when the states of the world follow the distribution σ_λ, the second-best profits $\pi^{sb}(0)$ are constant over all $\lambda \in [0,1]$. However, as the states grow further apart, having individualized contracts becomes more and more profitable and may eventually warrant providing the manager with information in spite of the additional costs incurred in doing so. The following proposition shows that linear mean-preserving contractions are sufficient to determine whether the owner prefers the manager to be perfectly informed or perfectly uninformed.

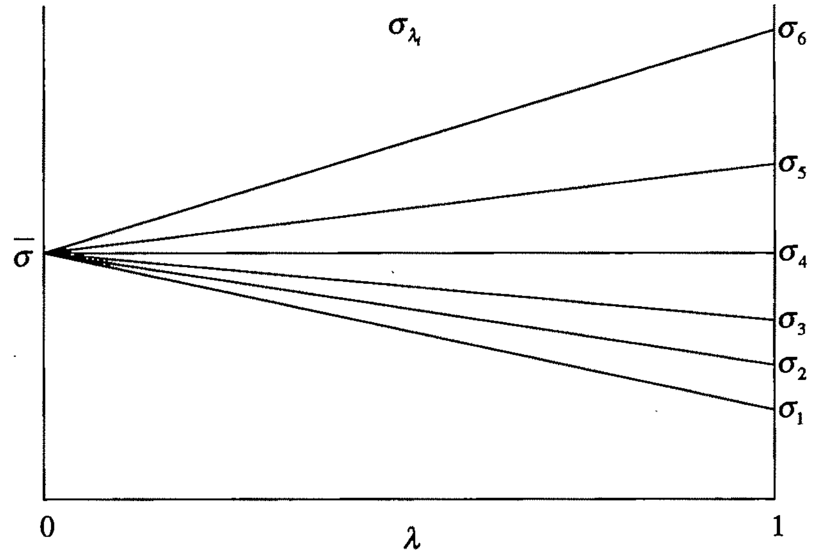
---

[9] There may be some confusion as to whether σ_1 is the random variable σ_λ with $\lambda = 1$, or whether σ_1 is an element in the support of σ. To avoid any confusion, we will explicitly label σ_1 as a random variable unless we are referring to the constant.

Proposition 5: Suppose that the owner prefers the manager's signal to be perfectly informative. Then, there exists a $\hat{\lambda} \in (0,1)$ such that the owner prefers the signal to be uninformative when the states of nature follow the lmpc σ_λ for all $\lambda < \hat{\lambda}$. If the owner prefers the signal to be perfectly uninformative, then she continues to prefer an uninformative signal when the states of nature follow the lmpc σ_λ for $\lambda \in [0,1]$.

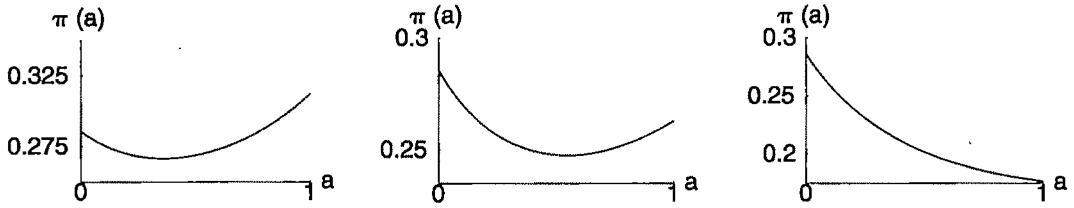
The proposition guarantees that given any distribution of the states of the world, the owner will always prefer to have an uninformed manager if all the possible states sufficiently resemble the mean $\bar{\sigma}$. Figure 3 plots an example of the owner's profits over *a* for various levels of λ.

FIGURE 2
Linear Mean-Preserving Contraction



We plot σ_{λ_i} for $i = 1,\ldots,6$. Note that σ_{λ_i} is equal to the original mean $\bar{\sigma}$ when $\lambda = 0$ and σ_{λ_i} is equal to the original support σ_i when $\lambda = 1$.

FIGURE 3
Convexity and Linear Mean-Preserving Contractions



The graphs plot a two-state example of the owner's expected profits over *a* for various levels of λ with $\sigma_1 = 1$, $\sigma_2 = 2.5$, and $f(\sigma_1) = f(\sigma_2) = \frac{1}{2}$. From left to right, the value of λ decreases from 1 to $\frac{3}{4}$ to $\frac{1}{4}$.

Collapsing the states into $\overline{\sigma}$ also affects the marginal profitability of offering a manager information. In the first panel of Figure 3, $\lambda = 1$ and the owner's profits are maximized at $a = 1$. As we decrease λ to $\frac{3}{4}$ in the second panel, the owner's profits are still increasing for sufficiently large a, but not enough to warrant setting $a = 1$. In the last panel of Figure 3, $\lambda = \frac{1}{4}$ and the owner's profits are always decreasing in a. In fact, the observation generalizes to any linear mean-preserving contraction: Proposition 3 and the proof of Proposition 5 directly imply that the principal's marginal profits $\pi^{sb'}(a)$ are increasing in λ. As such, the benefits to increasing the agent's private information are more pronounced when the states of nature are further spread out. For sufficiently small λ, the lmpc σ_λ will induce the owner's profits to be everywhere decreasing in a, i.e., she is worse off whenever the manager gains any additional informational advantage.

We have thus far considered changes in the support of σ that would cause the owner to change her informational preferences. A second possibility is that changing the probabilities associated with a fixed support could also cause the owner to switch her preferences over the choice of information system for the manager. We examine this next.

Example 1: Suppose there are three states, $\{\sigma_1,\sigma_2,\sigma_3\} = \{\sigma_1,\sigma_1 + k,\sigma_1 + 2k\}$, with respective probabilities $\left\{\frac{1-x}{2}, x, \frac{1-x}{2}\right\}$ and $x \in [0,1]$. In this case, as x tends to 1, the owner becomes convinced that the true state is $\sigma_1 + k$, whereas when x tends to 0, the owner believes that the true state of the world is either σ_1 or $\sigma_1 + 2k$, each with equal probability. To satisfy Assumption A1, we need $x \geq \sqrt{5} - 2 \approx 0.236$; Assumption A2 is trivially met. Now, from Proposition 4, the owner prefers $a = 1$ to $a = 0$ if and only if:

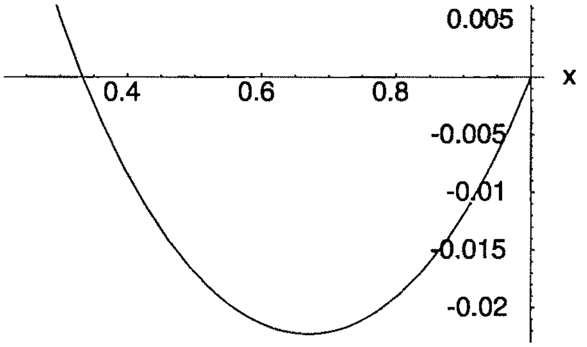
$$\frac{1-x}{2\sigma_1} + \frac{x}{\sigma_1 + k\dfrac{1+x}{2x}} + \frac{1-x}{2\left(\sigma_1 + k + \dfrac{2k}{1-x}\right)} - \frac{1}{\sigma_1 + k} \geq 0. \quad (5)$$

When $x = \frac{1}{3}$, i.e., the states are uniformly distributed, equality holds in (5) if $k = \frac{1+\sqrt{7}}{4}\sigma_1$. Figure 4 fixes $k = \frac{1+\sqrt{7}}{4}\sigma_1$ and plots the left-hand side of (5) as x varies from $\sqrt{5} - 2$ to 1.

As x tends to 1, the only possible state is σ_2, and communication becomes worthless as the first-best outcome can be replicated with a single contract. On the other hand, as x takes on smaller values, the owner becomes less and less certain of the true state and therefore the owner prefers to pay the manager rents in exchange for better information regarding σ.. In particular, when $k = \frac{1+\sqrt{7}}{4}\sigma_1$, for all values of $x \in [\frac{1}{3},1]$, the principal prefers to go with a single contract and avoid paying rents. For values of $x \in [\sqrt{5} - 2, \frac{1}{3}]$ there is sufficient uncertainty that the principal prefers to offer a menu of contracts and pay the manager the appropriate informational rents.

Example 1 illustrates the difficulty of using traditional stochastic orderings as a way to determine the owner's preferences over the manager's level of information superiority. In particular, note that both the value of an informed manager, as well as the relative value

FIGURE 4
Information Preferences and State Uncertainty



Where the graph is positive the owner prefers $a = 1$, and in the negative region, the owner prefers $a = 0$.

(relative to having an uninformed manager) are non-monotonic in x, despite the fact that a decrease in x represents a mean-preserving spread, a higher level of variance, and is riskier in the sense of second-order stochastic dominance. Therefore, a higher level of riskiness, *per se*, does not guarantee either that the absolute value or the relative value of having an informed manager is any greater. In the above example, note however that *once* the spread/riskiness is sufficiently large, the owner prefers to employ an informed manager. We now demonstrate that this finding can be generalized, under surprisingly mild conditions, to any arbitrary n-point distribution.

Proposition 6: For any arbitrary distribution satisfying Assumptions A1 and A2, there exists a sufficient mean-preserving spread of the probabilities $f(\cdot)$ under which the owner prefers the manager to be fully informed, provided that $\sigma_n \geq 2\sigma_1$.

In this section, we considered the effects of varying the manager's private information. We have shown that the owner's profits are initially decreasing and everywhere convex in the manager's private information. Hence, she will always prefer an uninformed manager to a poorly informed one, and will always prefer the manager to be either perfectly informed or perfectly uninformed. With regard to the latter, we have provided a full characterization of the circumstances under which the owner would prefer either outcome. In particular, sufficiently dispersed states of nature or probability levels over states imply a preference for a fully informed manager.

III. NOISY REVENUE

An important restriction in our analysis so far is the assumption that the outcome of the manager's effort was deterministic. As a consequence, we could restrict attention to menus consisting of revenue targets and attendant payoffs in order to control the manager. In this section, we analyze the impact of relaxing this restriction by considering settings where the manager's actions lead to stochastic revenue realizations.

Suppose the owner's revenue takes the form $\pi = \boldsymbol{\beta} \cdot \mathbf{e} + \varepsilon$ where the random variable ε has zero mean and finite, non-zero variance. The second-best contract analyzed in the previous section is then no longer feasible. One potential approach is to assume a specific functional form for the density of ε, and solve for the optimal contract for that particular density. We pursue an alternative strategy, that of restricting attention to menus of linear contracts. A key advantage of our approach is that the contracts we study are robust to *any* density of ε, i.e., they can be implemented even if the owner is unaware (or uncertain) of the precise distribution of the noise in revenues. A second important benefit is that by characterizing the optimal linear contracts, we are able to provide a direct comparison of our work to previous results on incentive strength and information. Menus of linear contracts have been extensively examined in the accounting literature (see Kirby et al. 1991; Reichelstein 1992), although such models typically use a continuum of types, whereas our model includes a finite type space.

The switch in settings from deterministic revenues to noisy revenues leads to a reduction in the owner's expected profits for the simple reason that menus of linear contracts are non-optimal in the deterministic setting. Linear contracts are costly in this environment because a type i agent can claim to be of type $i - 1$, and then select an effort level that generates greater payoffs than the principal's preferred effort level e_{i-1}. Note that if we had a continuum of types, then it is conceivable that the additional constraint associated with menus of linear contracts is costless, as the owner may already face the same constraint when solving for the optimal mechanism. Indeed, Rogerson (1987) demonstrates that, with a continuum of types and signals, one can restrict attention to menus of linear contracts, provided the payments to the manager in the deterministic revenue setting are convex in output. In our framework, however, even in a two-type setting where the payments to the manager are trivially convex in type, menus of linear contracts impose an additional cost to the owner, irrespective of whether the owner tries to implement the original second-best effort levels, or the optimal effort levels given a menu of linear contracts. In general, the optimal linear menu of contracts will implement different effort levels than the optimal second-best contract, even in the absence of output uncertainty.

We consider linear contracts that consist of a fixed payment, x_l, and a share of firm revenues y_l. If the manager has received the signal $s = s_i$, and subsequently chooses the contract pair $\{x_l, y_l\}$, then he solves:

$$\max_{\mathbf{e}} \; y_l(\boldsymbol{\beta} \cdot \mathbf{e}) + x_l - \frac{\hat{\sigma}_i}{2} \sum_{j=1}^{k} e_j^2 .$$

As such, we find that a type i manager facing an l linear contract will put forth effort:

$$e_{i,j} = \frac{y_l \beta_j}{\hat{\sigma}_i},$$

which will generate a total expected revenue of $r_{i,l} = (\boldsymbol{\beta} \cdot \boldsymbol{\beta})\, y_l / \hat{\sigma}_i$. When the owner offers a menu of linear contracts, she thus solves:

$$\max_{\mathbf{x},\mathbf{y}} \sum_{i=1}^{n} f(\sigma_i)((1 - y_i) r_{i,i} - x_i) \tag{P2}$$

such that:

$$y_i r_{i,i} + x_i - c_i(r_{i,i}) \geq 0 \quad \forall i$$
$$y_i r_{i,i} + x_i - c_i(r_{i,i}) \geq y_l r_{i,l} + x_l - c_i(r_{i,l}) \quad \forall i,l.$$

The first constraint above is identical to that found in the deterministic setting. The second constraint ensures that the manager always truthfully reveals his signal, even though he can subsequently choose from a continuum of effort levels.[10] As in the deterministic setting, we must make two additional assumptions to guarantee that the solution to P2 is well behaved. We will again assume A1 and we need to slightly alter A2 to accommodate the stochastic setting. We will assume that the following modification of A2 is satisfied:

A2*: The ratio σ_i/σ_{i-1} is increasing in i.

Assumption A2* replaces A2 from the deterministic setting and serves to maintain the monotonicity of the effort schedule with menus of linear contracts. Note that A2* holds for a large number of situations, including the case where the support of σ follows a geometric series. Assumption A2* is sufficient, but not necessary; the "natural" setting where σ has an equally spaced support and is uniformly distributed continues to be acceptable here, in spite of violating A2*. The following proposition gives the solution to P2.

Proposition 7: Assuming A1 and A2*, the solution to P2 is given by:

$$y_i = \frac{1}{\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1} \tag{6}$$

$$x_i = x_{i+1} + \frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{2\hat{\sigma}_i}(y_{i+1}^2 - y_i^2) \tag{7}$$

$$x_n = -\frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{2\hat{\sigma}_n} y_n^2 \tag{8}$$

The solution to P2 is similar to that of P1: managers with better signals (smaller indices) put forth more effort than others and are also entitled to larger rents. The worst type of manager earns no rents, and the best type puts forth the same level of effort as in the first-best setting (i.e., when the signal is observed by the owner). However, the optimal menu of linear contracts distorts the manager's effort beyond the second-best contract found in the deterministic setting. To see this, let $e_{i,j}^l$ be the effort put forth on task j by a manager with signal $s = s_i$ facing a menu of linear contracts, and let $e_{i,j}^d$ be the same effort when the manager faces the deterministic menu found in the last section. We have:

---

[10] Recall that in the deterministic setting, the manager did not need to communicate to the owner; it was sufficient for him to choose a revenue target that, in effect, communicated the manager's signal to the owner. With noisy outcomes, the manager cannot communicate his signal via revenue, because he can no longer ensure that a specific target level of revenue is achieved. Instead, the manager communicates his information to the owner by selecting a linear contract from the menu before selecting his effort levels.

$$\begin{aligned} e^{l}_{ij} &= \frac{1}{\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1} \cdot \frac{\beta_j}{\hat{\sigma}_i} \\ &\leq \frac{1}{\frac{F(\sigma_i)}{f(\sigma_i)}\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1} \cdot \frac{\beta_j}{\hat{\sigma}_{i-1}} = e^{d}_{ij} \\ &\leq \frac{\beta_j}{\hat{\sigma}_i} = e^{fb}_{ij}. \end{aligned}$$

The owner needs to distort the effort levels beyond those found in the deterministic second-best contract in order for the contract to satisfy the agent's incentive compatibility constraint for all possible densities of ε. In spite of the difference in effort levels between the deterministic and stochastic settings, our next result shows that the nature of the relation between a and the owner's profits is unaltered.

Proposition 8: Using the optimal menu of linear contracts, the owner's second-best profits are decreasing in a at $a = 0$, and are strictly convex over the entire range of a.

An important implication of Proposition 8 is that all of our results on the value of informing the manager continue to hold in the stochastic framework. Moreover, with menus of linear contracts, the owner still prefers an uninformed manager when the states of the world are relatively close to one another and prefers an informed manager when the states are spaced widely apart, as the next result demonstrates.

Proposition 9: In the stochastic setting, if the owner can choose the level of information conveyed in s, she will make s perfectly informative when:

$$\frac{1}{\bar{\sigma}} \leq \mathrm{E}\left[\frac{\sigma_i^{-1}}{1 + \left(\frac{F(\sigma_i)}{f(\sigma_i) - 1}\right)\left(\frac{\sigma_i}{\sigma_{i-1}} - 1\right)}\right]$$

or, equivalently,

$$\frac{1}{\bar{\sigma}} \leq \mathrm{E}\left[\frac{y_i(1)}{\sigma_i}\right],$$

and perfectly uninformative otherwise.

The underlying intuition behind the owner's preferences over a remains entirely unchanged between the deterministic and stochastic settings. As such, all of our results on linear mean-preserving contractions continue to hold. In particular, for any given random variable σ, there always exists a linear mean-preserving contraction under which the principal will prefer $a = 0$ to $a = 1$. For the example considered in Corollary 1, we find that when the support of σ is given by $\{\sigma_1, \rho\sigma_1, \ldots, \rho^{n-1}\sigma_1\}$, the principal will prefer $a = 1$ to a

= 0 whenever $\rho \geq 3$. Note that in the stochastic setting, the support of σ needs to be further spread out than in the deterministic setting for information to be valuable. This follows from the fact that in the stochastic setting, the principal faces an additional constraint in motivating the agent to select the correct effort level in addition to reporting truthfully. It can also be shown that Example 1 carries over to the stochastic setting, as does the subsequent existence result proving the owner's preference for an informed manager under extreme mean-preserving spreads.

In addition to being robust to any type of white noise, the use of menus of linear contracts allows us to explore the effect of information on incentives. Baker and Jorgensen (2003) is one of the first papers to explore the relationship between incentives and private information. In their model, a risk-averse manager's private information cannot be communicated to the owner. Instead, given the level of information contained in the manager's signal, the owner offers the manager a single linear contract. As the slope of the linear contract increases, the manager's effort level converges to the level the owner would have chosen had she known the manager's private information. The owner can recoup the cost of offering strong incentives via an upfront payment from the manager to the owner. However, the owner does not offer the manager perfect incentives because doing so would require the owner to pay the manager a large risk premium, which is a deadweight loss that cannot be recouped. Baker and Jorgensen (2003) find that as the manager's signal becomes more informative, the marginal return to good decision making increases, which in turn induces the owner to offer stronger incentives to motivate the manager to make better decisions. Similar to Baker and Jorgensen's (2003) model, as a, the informativeness of the signal, increases in our setting, the value of making informed decisions increases. However, because we allow for communication via menus of contracts, the principal has more control over the decisions, and hence the extent to which decisions are informed is not necessarily increasing in incentive strength. The following result shows that in our model, as the manager's signal becomes more informative, it is in fact optimal for the principal to provide him with *decreased* incentives.

Proposition 10: The slope of each contract from the optimal menu of linear contracts is weakly decreasing in a. The average slope is strictly decreasing in a for $a > 0$.

The intuition behind this proposition rests in the IC constraint of P2. As a increases, i.e., the manager's information becomes more precise, the expected states of the world $\{\hat{\sigma}_1,...,\hat{\sigma}_n\}$ grow further apart, which makes the linear contract (x_{i+1}, y_{i+1}) appear more attractive to a manager with signal i. In order to keep a type i manager truthful, the owner must pay additional informational rents. However, the owner minimizes such rents by further distorting the effort level ultimately chosen by the manager via decreased incentives.

The above finding shows that the owner will always choose to distort the manager's effort levels in response to increases in a. However, this does not imply that the manager's rents are increasing in a. Instead, in response to an increase in a, it is possible that the owner distorts the manager's choice until his informational rents actually decrease. In the next section, we consider the manager's preferences over a; in addition, we solve for the welfare-maximizing level of private information and compare both to the owner's preferred choice of a.

IV. PREFERENCES

Thus far, we have considered the choice of information systems only from the standpoint of the firm's owner. As discussed in the introduction, it is often the case that the manager can influence his relative informational superiority over the owner. For example, the manager can choose to acquire better local knowledge, expertise, or additional data regarding specific markets, products, or production processes. Such data may be either unavailable to, or less interpretable by, the firm's owners, thereby increasing the manager's informational advantage. In this section, we begin by examining the manager's preferences over the level of *ex ante* information quality. While it may seem obvious that a manager would always choose to maximize his informational superiority over the owner, we show that this is in fact not the case.

To keep the analysis tractable, in this section we restrict ourselves to only two equally likely states ($n = 2$). We begin with the deterministic setting and discuss the stochastic setting later. Suppose that $\sigma_2 = \rho\sigma_1$. From the proof of Corollary 1, we can deduce that the owner prefers $a = 1$ if and only if $\rho \geq 2$. From a social welfare perspective, the optimal value of a is that which maximizes the owner's revenues minus the manager's cost of effort. Formally, if we let $W(a)$ denote welfare, then:

$$\begin{aligned} W(a) &= \sum_{i=1}^{n} f(\sigma_i)\left(r_i(a) - \frac{\hat{\sigma}_i(a) r_i^2(a)}{2(\beta \cdot \beta)}\right) \\ &= \nu \frac{1}{2\sigma_1}\left(\frac{1}{\rho(1-a)+1+a} + \frac{2}{3a(\rho-1)+\rho+1} \right. \\ &\quad \left. - \frac{a(\rho-1)+\rho+1}{(3a(\rho-1)+\rho+1)^2}\right), \end{aligned}$$

where ν is a positive constant. The following lemma shows that in the region $a \in [0,1]$, the function $W(a)$ is maximized at either $a = 0$ or $a = 1$.

Lemma 2: Social welfare $W(a)$ is maximized at $a = 1$ if $\rho \geq 1.25$; if $\rho \leq 1.25$, then $W(a)$ is maximized at $a = 0$.

Similar to the principal's preferences, welfare is maximized at the boundaries of a; yet, the two do not always coincide. In particular, for values of $\rho \in [1.25,2]$, the owner prefers the manager to be perfectly uninformed, whereas the welfare-maximizing solution calls for the manager to be perfectly informed. To understand the divergence, recall that the only difference between overall welfare and the owner's payoffs are the informational rents paid to the manager. The manager's rents can be expressed as $R(a)$, where:

$$\begin{aligned} R(a) &= \sum_{i=1}^{n} f(\sigma_i)(p_i - c_i(r_i(a))) \\ &= \eta_1 \frac{a(\sigma_2 - \sigma_1)}{2(a\sigma_1 + (1-a)\bar{\sigma} + 2a(\sigma_2 - \sigma_1))^2} \\ &= \eta_2 \frac{a}{\sigma_1((3a(\rho-1)+\rho+1)^2}. \end{aligned}$$

Here, η_1 and η_2 are positive constants. The following result identifies the manager's preferred level of private information quality, i.e., the one that maximizes his informational rents.

Proposition 11: The manager prefers $a = 1$ for $\rho \leq 2$; for $\rho > 2$, the manager prefers the interior solution $a = \dfrac{1 + \rho}{3(\rho - 1)} \in (0,1)$.

A key implication of the above proposition is that for sufficiently large ρ, the manager wants to limit his amount of private information relative to the owner. In other words, even though the manager could costlessly choose to be perfectly informed, he prefers not to, and settles for an interior level of information quality. To see why, note that for sufficiently large a, the incentive compatibility constraint weakens; the manager realizes that submitting the wrong message results in heavily distorted revenue targets, so much so that he begins to develop a preference for telling the truth. As a consequence, the manager's informational rents start to decrease in a beyond a threshold level, say $\hat{a}$. Moreover, the manager's preferred level of (interior) information quality $\hat{a}$, decreases monotonically in ρ for values of $\rho > 2$, i.e., as the disparity in the set of possible state realizations grows. Because an uninformed manager earns zero rents, we find that $\hat{a} > 0$, however; in the limit, as $\rho \rightarrow \infty$, the lower bound on $\hat{a}$ is $1/3$.

A second important implication of the lemma is that the manager's preferred level of a *never* coincides with the owner's preferences. At the precise cutoff ($\rho = 2$) where the principal's preferences switch from $a = 0$ to the welfare-maximizing $a = 1$, the manager's preferences switch from $a = 1$ to an interior value $\hat{a} < 1$.

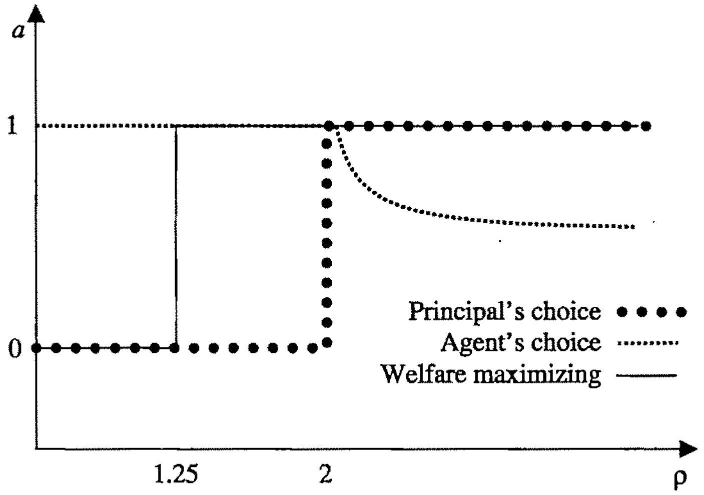
A final implication is illustrated in Figure 5, which shows the owner's choice of a and juxtaposes it over the welfare-maximizing solution and the manager's choice. Note that the welfare-maximizing level of a agrees with the owner's choice when ρ takes on extreme values, in particular when either $\rho \geq 2$ or $\rho \leq 1.25$. For values of ρ in the middle, i.e., $\rho \in [1.25,2]$, the socially optimal level of a coincides with the manager's preferred choice. Put differently, if the owner will not choose the welfare-maximizing level of information quality, the manager will. As a consequence, because the contracts considered take a as fixed, total welfare will increase if the principal delegates the choice of *ex ante* information system quality to the manager when the spread between the states takes on intermediate values.

Finally, we address the issue of delegation in the context of uncertain revenues and linear contracts. The following lemma recreates the above analysis for the case of stochastic revenues and optimal menus of linear contracts:

Lemma 3: With the optimal menu of linear contracts, the owner prefers $a = 0$ if $\rho \leq (1 + \sqrt{2})$, and $a = 1$ otherwise. The manager prefers $a = 1$ for $\rho \leq \sqrt{3}$ and $a = (2 - \sqrt{3})\dfrac{\rho + 1}{\rho - 1}$ otherwise. Overall welfare is maximized at $a = 0$ for $\rho \leq \frac{1}{2}(1 + \sqrt{5})$; otherwise, $a = 1$ maximizes welfare.

As in the deterministic setting, with stochastic revenues and optimal menus of linear contracts, the manager reacts to sufficiently high levels of ρ (i.e., very disparate types) by deliberately restricting his quality of private information. In particular, for values of $\rho > \sqrt{3}$, the manager chooses $a = (2 - \sqrt{3})\dfrac{\rho + 1}{\rho - 1} < 1$. The manager's choice of a is

FIGURE 5
Informational Preferences



The socially optimal choice of a, in addition to the owner's and the manager's preferred choices, in the deterministic setting.

strictly decreasing in ρ and converges to $a = 2 - \sqrt{3}$ as $\rho \rightarrow \infty$. A second similarity with the deterministic setting is the divergence of the manager and owner's *ex ante* informational preferences. It is important to note that this result does not follow from the fact that the manager wishes to maximize rents and the owner seeks to minimize them. Instead, as *a* increases, the firm becomes (potentially) more profitable, as seen in Proposition 1, but in addition, larger distortions are introduced to limit the manager's rents. As such, there is no guarantee that increasing *a* does not increase both the owner's profits and manager's rents. Our result that the manager and owner disagree unambiguously is thus quite striking.

A key difference from the deterministic setting is that, with stochastic revenues, there exists a region where neither the manager nor the owner wants to choose the welfare-maximizing level of private information quality. In particular, in the region $\rho \in [\sqrt{3}, 1 + \sqrt{2}]$, overall welfare is maximized by setting *a* to 1; however, the owner prefers $a = 0$ while the manager prefers an interior level of *a*. In this region, it can be shown that the agent's choice of *a* yields greater joint payoffs than the principal's choice of *a*. Similar to the deterministic setting, overall welfare will thus increase if the principal can commit to allowing the agent to choose *a* before the two sign a contract. Note that in both the deterministic and stochastic settings, this result occurs for intermediate values of ρ. For extreme values of ρ (i.e., when the states of the world are either similar or sufficiently far apart), the principal prefers the welfare-maximizing level of *a*, and would not consider delegating the choice of information system quality to the manager.

V. CONCLUSION

In a setting where an owner delegates decisions to a manager, having a better-informed manager is not always preferred when the information is privately observed by the manager. In our model, we show that the owner's payoffs are initially decreasing and everywhere strictly convex in the correlation between the manager's signal and the true state of the world. As such, the owner will always prefer to deal with an uniformed manager to a poorly informed one, and will always opt for either an uninformed or a perfectly informed manager over a partially informed one. The owner's preference among these extreme options is a function of the distribution of nature, in a manner made precise in our analysis.

The above results have implications for the design of managerial accounting systems where an owner cannot interpret incoming information but can control the flow of information to her managers. We have shown that when the optimal effort level does not vary significantly, there is little use to informing the manager because doing so will cost the owner informational rents and only bring about small gains from more informed decision making.

Our analysis also has implications in settings where an owner cannot control the flow of private information to her managers. In such a setting, we show that it is optimal for the owner to offer the manager weaker incentives as the manager's level of private information increases. Even more surprisingly, we show that a manager who has the option to increase his level of private information relative to the owner will often deliberately choose to curtail this advantage, because his informational rents are not necessarily increasing in the quality of his private information. Moreover, we demonstrate that the manager's preferred level of information quality often more closely matches the welfare-maximizing level of information.

Our findings hold in settings where the manager is able to achieve set revenue targets in a deterministic fashion, as well as when revenues are a stochastic function of the manager's effort levels. The majority of our results were generated with relatively few parametric restrictions. We allowed for an arbitrary number of managerial tasks and associated effort levels, an arbitrary number of informational signals, and (with one restriction) any density function on those signals. One key aspect of our analysis, however, required additional restrictions. In particular, for reasons of tractability, we derived the informational preferences of the manager, as well as the results on disagreement over overall welfare, in a two-state setting. We believe that these results hold in greater generality though, and have verified them via simulations for the $n = 3$ and $n = 4$ cases. It would be desirable to prove these results for the general $n > 2$ case.

When analyzing the owner's preferences, we introduced the notion of linear mean-preserving contractions. Using this construct, we were always able to distort the states of the world σ until the owner preferred an uninformed manager to an informed one. It is infeasible to create a similar linear distortion in the opposite direction to favor an informed manager without potentially including negative supports in the distribution of σ. However, it may be possible to construct general nonlinear expansions that would achieve this objective. With regard to the impact of varying densities, we demonstrated, under mild conditions, the existence of a mean-preserving spread that led to the owner's preference for an informed manager. We conjecture that the following, sharper characterization may be possible: for an arbitrary discrete density, *if* the owner prefers an informed manager, *then* he will continue to prefer an informed manager for any mean-preserving spread of that density.

Finally, while we have framed our analysis in terms of internal accounting systems, our results apply to any setting where incentive contracts are written among parties with asymmetric information. An especially compelling extension of our work is to buyer-supplier contracting, where the former, because of proximity to customers, typically possesses superior demand forecast information (Cachon and Lariviere 2001). An important difference is that the demand information is communicated to a supplier who then makes production decisions that affect the interests of both parties. In future research, it would be of great interest to analyze such a model in full generality.

APPENDIX

Proof of Lemma 1

Suppose that the unconditional means of s and σ are given by $\bar{s}$ and $\bar{\sigma}$, respectively. Denote the joint density of s and σ by $h(s_i, \sigma_l)$. We have:

$$\begin{aligned}
\operatorname{cov}(s, \sigma) &= \sum_{i=1}^{n} \sum_{l=1}^{n} (s_i - \bar{s})(\sigma_l - \bar{\sigma})h(s_i, \sigma_l) \\
&= \sum_{i=1}^{n} \sum_{l=1}^{n} (s_i - \bar{s})(\sigma_l - \bar{\sigma})f(s_i)\Pr[\sigma = \sigma_l \mid s = s_i, a] \\
&= \sum_{i=1}^{n} (s_i - \bar{s})f(s_i)\left(a(\sigma_i - \bar{\sigma}) + (1 - a)\sum_{l=1}^{n} (\sigma_l - \bar{\sigma})f(\sigma_l)\right) \\
&= a\sum_{i=1}^{n} (s_i - \bar{s})(\sigma_i - \bar{\sigma})f(s_i)
\end{aligned}$$

Therefore we can write:

$$\operatorname{corr}(s, \sigma) = a\frac{\sum_{i=1}^{n} (s_i - \bar{s})(\sigma_i - \bar{\sigma})f(s_i)}{\sqrt{\sum_{i=1}^{n} (s_i - \bar{s})^2\, f(s_i)}\,\sqrt{\sum_{i=1}^{n} (\sigma_i - \bar{\sigma})^2\, f(\sigma_i)}}.$$

Since s is simply a signal, we can choose the actual signals s_i without loss of generality; in particular, we can set $s_i = \sigma_i$, in which case $\operatorname{corr}(s, \sigma) = a$.

Proof of Proposition 1

Differentiating the expression for the owner's profit with respect to a yields:

$$\pi^{fb\prime}(a) = -\frac{(\beta \cdot \beta)}{2} \sum_{i=1}^{n} f(\sigma_i)\, \frac{(\sigma_i - \bar{\sigma})}{\hat{\sigma}_i^2}.$$

As such, note that:

$$\pi^{fb\prime}(0) = -\frac{(\beta \cdot \beta)}{2\bar{\sigma}^2} \sum_{i=1}^{n} f(\sigma_i)(\sigma_i - \bar{\sigma}) = 0.$$

Moreover, we have:

$$\pi^{fb\prime\prime}(a) = (\beta \cdot \beta) \sum_{i=1}^{n} f(\sigma_i)\, \frac{(\sigma_i - \bar{\sigma})^2}{\hat{\sigma}_i^3} > 0.$$

Proof of Proposition 2

We shall omit the dependence of r, p, and σ upon a for this proof. Note that (2) implies that if (1) holds for $i = n$, then it must also hold for $i < n$, as:

$$p_i - c_i(r_i) \geq p_n - c_i(r_n) > p_n - c_n(r_n) \geq 0.$$

The first inequality above holds by the IC constraint, the second inequality holds because $\hat{\sigma}_i < \hat{\sigma}_n$ for $i < n$, and the last inequality holds by the IR constraint, and is met with equality in the case of $i = n$. As such, we can, with no loss of generality, set $p_n = c_n(r_n)$. We will only consider the local upward incentive compatibility constraints of the form:

$$p_i - c_i(r_i) \geq p_{i+1} - c_i(r_{i+1})$$

and will verify that the optimal solution to this problem satisfies the entire set of constraints (2).

If the local upward IC constraints bind, then the optimal contract satisfies:

$$p_i = p_{i+1} + \frac{\hat{\sigma}_i}{2(\beta \cdot \beta)}(r_i^2 - r_{i+1}^2).$$

We can express this equivalently as follows:

$$p_i = p_n + \sum_{l=i}^{n-1} \frac{\hat{\sigma}_l}{2(\beta \cdot \beta)}(r_l^2 - r_{l+1}^2). \tag{9}$$

The owner's problem can thus be re-stated as:

$$\max_{\mathbf{r}} \sum_{i=1}^{n} f(\sigma_i)\left(r_i - \frac{\hat{\sigma}_n r_n^2}{2(\beta \cdot \beta)} - \sum_{l=i}^{n-1} \frac{\hat{\sigma}_l}{2(\beta \cdot \beta)}(r_l^2 - r_{l+1}^2)\right). \tag{10}$$

Abusing notation, let $F(\sigma_0) = 0$. Then, we can rewrite the owner's objective as:

$$\max_{\mathbf{r}} \frac{1}{\beta \cdot \beta} \sum_{i=1}^{n}\left[f(\sigma_i)(\beta \cdot \beta)r_i - F(\sigma_i)\frac{\hat{\sigma}_i}{2}r_i^2 + F(\sigma_{i-1})\frac{\hat{\sigma}_{i-1}}{2}r_i^2\right].$$

Since our unconstrained objective is quadratic and concave in each r_i, the unique solution to our problem can be found via the first order condition, yielding:

$$r_i = \frac{(\beta \cdot \beta)}{\frac{F(\sigma_i)}{f(\sigma_i)}(\hat{\sigma}_i - \hat{\sigma}_{i-1}) + \hat{\sigma}_{i-1}}.$$

We claim that r_i is decreasing in i. To prove this claim, it is sufficient to show that the denominator is increasing in i. Recall that A1 assumes that $F(\sigma_i)/f(\sigma_i)$ is increasing in i, and $\hat{\sigma}_{i-1}$ is increasing in i by construction. Therefore, it suffices to show that $\hat{\sigma}_i - \hat{\sigma}_{i-1}$ is increasing in i. The following statements are equivalent:

$$\hat{\sigma}_i - \hat{\sigma}_{i-1} \geq \hat{\sigma}_{i-1} - \hat{\sigma}_{i-2}$$

$$a\sigma_i + (1-a)\bar{\sigma} - a\sigma_{i-1} - (1-a)\bar{\sigma} \geq a\sigma_{i-1} + (1-a)\bar{\sigma} - a\sigma_{i-2} - (1-a)\bar{\sigma}$$

$$a\sigma_i - a\sigma_{i-1} \geq a\sigma_{i-1} - a\sigma_{i-2}.$$

The last line above holds by A2. By construction, the solution satisfies local upward incentive compatibility. The following equivalent statements verify global upward incentive compatibility, i.e., that type i would never want to take a type $i + l$ contract.

$$\begin{aligned} &p_i - c_i(r_i) \geq p_{i+l} - c_i(r_{i+l}) \\ &p_i \geq p_{i+l} - c_i(r_{i+l}) + c_i(r_i) \end{aligned} \tag{11}$$

$$\sum_{t=i}^{i+l-1} \frac{\hat{\sigma}_t}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} (r_t^2 - r_{t+1}^2) \geq \frac{\hat{\sigma}_i}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} (r_i^2 - r_{i+l}^2). \tag{12}$$

To go from (11) to (12), we use (9). The inequality in (12) holds since r_i^2 is decreasing in i whereas $\hat{\sigma}_i$ is increasing in i.

Finally, we need to verify that type i would not want to mimic some type $i - l$. The following equivalent statements verify global downward incentive compatibility:

$$\begin{aligned} p_i = c_i(r_i) &\geq p_{i-l} - c_i(r_{i-l}) \\ p_i &\geq p_{i-l} - c_i(r_{i-l}) + c_i(r_i) \\ 0 &\geq \sum_{t=i-l}^{i-1} \frac{\hat{\sigma}_t}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} (r_t^2 - r_{t+1}^2) + \frac{\hat{\sigma}_i}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} (r_i^2 - r_{i-l}^2). \end{aligned}$$

The final inequality holds again by the monotonicity of $\hat{\sigma}_i^2$ and r_i^2 in i.

Proof of Proposition 3

Label the owner's unconstrained objective function as $M(a, \mathbf{r})$. Again, label $F(\sigma_0) = 0$. Ignoring the scaling of $(\boldsymbol{\beta} \cdot \boldsymbol{\beta})$, we can write:

$$M(a, \mathbf{r}) = \sum_{i=1}^{n} \left[f(\sigma_i)(\boldsymbol{\beta} \cdot \boldsymbol{\beta})r_i - F(\sigma_i)\frac{\hat{\sigma}_i}{2} r_i^2 + F(\sigma_{i-1})\frac{\hat{\sigma}_{i-1}}{2} r_i^2 \right].$$

Let $V(a) = \max_{\mathbf{r}} M(a, \mathbf{r})$. The envelope theorem yields:

$$\begin{aligned} V'(a) = M_a(a, \mathbf{r}) &= \sum_{i=1}^{n} \left[-F(\sigma_i)\frac{(\sigma_i - \bar{\sigma})}{2} r_i^2 + F(\sigma_{i-1})\frac{(\sigma_{i-1} - \bar{\sigma})}{2} r_i^2 \right] \\ &= \sum_{i=1}^{n} \left[F(\sigma_{i-1})\frac{(\sigma_{i-1} - \sigma_i)}{2} r_i^2 - f(\sigma_i)\frac{(\sigma_i - \bar{\sigma})}{2} r_i^2 \right]. \end{aligned}$$

If $a = 0$, then $r_i = r$ for $i = 1,\ldots,n$; hence, we can write:

$$V'(0) = r^2 \sum_{i=2}^{n} F(\sigma_{i-1})\frac{(\sigma_{i-1} - \sigma_i)}{2} < 0,$$

where the last inequality holds by the monotonicity of σ_i in i.

To show that profits are strictly convex over a, we first prove a useful mathematical fact:

Claim: Suppose $G(a) = \max_{\mathbf{z}} g(a, \mathbf{z})$, and $\mathbf{z}(a) \in \text{argmax } g(a, \mathbf{z})$. Then if $g_{aa}(a, \mathbf{z}(a)) \geq 0$, G is convex in a. Moreover, if either $g_{aa}(a, \mathbf{z}(a)) > 0$ or $\mathbf{z}(a)$ is the unique maximizer of $g(a, \mathbf{z})$, then G is strictly convex in a.

For arbitrary $\alpha \in [0,1]$, let $\alpha \underline{a} + (1 - \alpha)\bar{a} = \tilde{a}$. Now, we have:

$$\begin{aligned} G(\tilde{a}) &= g(\tilde{a}, \mathbf{z}(\tilde{a})) \leq \alpha g(\underline{a}, \mathbf{z}(\tilde{a})) + (1 - \alpha) g(\bar{a}, \mathbf{z}(\tilde{a})) \\ &\leq \alpha g(\underline{a}, \mathbf{z}(\underline{a})) + (1 - \alpha) g(\bar{a}, \mathbf{z}(\bar{a})) = \alpha G(\underline{a}) + (1 - \alpha) G(\bar{a}). \end{aligned}$$

The first inequality holds by the convexity of g, and the second by the optimality of $\mathbf{z}(a)$. Clearly if either or both inequalities are strict, then G is strictly convex.

Now, returning to our main result, note that the unconstrained objective function M is linear in a, i.e., $M_{aa}(a, \mathbf{r}) = 0$. Since M is quadratic in each element of $\mathbf{r}$, there is a unique maximizer of M, hence $V''(a) > 0$ by the above claim.

Proof of Proposition 4

By Proposition 2 we have $\pi^{sb}(0) = (\boldsymbol{\beta} \cdot \boldsymbol{\beta}) \frac{1}{2\bar{\sigma}}$. Now, the principal prefers the agent to be informed if any of these equivalent statements hold:

$$\begin{aligned} &\pi^{sb}(0) \leq \pi^{sb}(1) \\ &(\boldsymbol{\beta} \cdot \boldsymbol{\beta}) \frac{1}{2\bar{\sigma}} \leq \sum_{i=1}^{n} f(s_i) r_i - \frac{1}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} \sum_{i=1}^{n} r_i^2 [\sigma_i F(\sigma_i) - \sigma_{i-1} F(\sigma_{i-1})] \\ &(\boldsymbol{\beta} \cdot \boldsymbol{\beta}) \frac{1}{2\bar{\sigma}} \leq \sum_{i=1}^{n} f(s_i) r_i - \frac{1}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} \sum_{i=1}^{n} r_i^2 \left[(\sigma_i - \sigma_{i-1}) \frac{F(\sigma_i)}{f(\sigma_i)} + \sigma_{i-1} \right] f(\sigma_i). \end{aligned} \tag{13}$$

By Proposition 2, we can restate (13) as:

$$\begin{aligned} &(\boldsymbol{\beta} \cdot \boldsymbol{\beta}) \frac{1}{2\bar{\sigma}} \leq \sum_{i=1}^{n} f(s_i) r_i - \frac{1}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} \sum_{i=1}^{n} r_i^2 \left[\frac{\boldsymbol{\beta} \cdot \boldsymbol{\beta}}{r_i} \right] f(\sigma_i) \\ &(\boldsymbol{\beta} \cdot \boldsymbol{\beta}) \frac{1}{2\bar{\sigma}} \leq \sum_{i=1}^{n} f(s_i) r_i - \frac{1}{2} \sum_{i=1}^{n} r_i f(\sigma_i) = \frac{1}{2} \sum_{i=1}^{n} r_i f(\sigma_i). \end{aligned}$$

Proof of Corollary 1

Proposition 2 implies that the owner prefers $a = 0$ if any of the following equivalent statements hold:

$$\begin{aligned} \pi^{sb}(0) &\geq \pi^{sb}(1) \\ \frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta}) n(\rho - 1)}{2\sigma_1(\rho^n - 1)} &\geq \frac{\boldsymbol{\beta} \cdot \boldsymbol{\beta}}{2n\sigma_1} \sum_{i=2}^{n} \frac{1}{\rho^{i-2}(1 + (\rho - 1)i)} \\ \frac{n^2(\rho - 1)}{(\rho^n - 1)} &\geq \sum_{i=2}^{n} \frac{1}{\rho^{i-2}(1 + (\rho - 1)i)} \\ \frac{n^2}{1 + \rho + \rho^2 + \ldots + \rho^{n-1}} &\geq \sum_{i=2}^{n} \frac{1}{\rho^{i-2}(1 + (\rho - 1)i)}. \end{aligned} \tag{14}$$

When $\rho = 1$, it is clear that equality holds in (14). Note that:

$$\lim_{\rho \to 1} \frac{\partial}{\partial \rho} \left[\frac{n^2}{1 + \rho + \rho^2 + \ldots + \rho^{n-1}} \right] = \frac{n(1 - n)}{2}$$

$$\lim_{\rho \to 1} \frac{\partial}{\partial \rho} \sum_{i=2}^{n} \frac{1}{\rho^{i-2}(1 + (\rho - 1)i)} = n(1 - n).$$

By continuity, we find that for sufficiently small ρ the owner will prefer $a = 0$. There also exists $\tilde{\rho}(n)$ such that, if $\rho > \tilde{\rho}(n)$, the owner prefers $a = 1$. We will now show that $\tilde{\rho} \leq 2$. Recall from (14) that the owner prefers $a = 1$ if and only if:

$$\frac{n^2}{1 + \rho + \rho^2 + \ldots + \rho^{n-1}} \leq \sum_{i=1}^{n} \frac{1}{\rho^{i-2}(1 + (\rho - 1)i)}. \tag{15}$$

A simple calculation reveals that the inequality in (15) holds for $n = 2$ and $\rho = 2$. To show that (15) holds for $n \geq 3$, note that the right-hand side is clearly increasing in n. It is therefore sufficient to show that the left-hand side is decreasing in n for $n \geq 3$. To see this, note that we can rearrange the left hand side of (15):

$$\frac{n^2}{1 + \rho + \rho^2 + \ldots + \rho^{n-1}} = \frac{n^2(\rho - 1)}{\rho^n - 1}.$$

Moreover:

$$\frac{\partial}{\partial n} \left[\frac{n^2(\rho - 1)}{\rho^n - 1} \right] = -\frac{(\rho - 1)n(2 - 2\rho^n + n\log[\rho]\rho^n)}{(1 - \rho^n)^2}. \tag{16}$$

A sufficient condition for the right-hand side of (16) to be negative is $n\log[\rho] - 2 \geq 0$, and since $3\log[2] > 2$, we find that the right-hand side of (16) is negative for $n \geq 3$ when $\rho = 2$.

Proof of Proposition 5

Define $\pi(a, \lambda)$ as second-best profits when the states of nature follow the random variable σ_λ. We will show that $\pi^{SB}(a, \lambda) = \pi^{SM}(\lambda a)$. From (10), second-best profits are a function of $\hat{\sigma}$ and not σ; therefore, it suffices to show that $\hat{\sigma}_\lambda(a) = \hat{\sigma}(\lambda a)$. To this end, we have:

$$\begin{aligned}
\hat{\sigma}_{\lambda_i}(a) &= a\sigma_{\lambda_i} + (1 - a)\overline{\sigma} \\
&= a(\lambda\sigma_i + (1 - \lambda)\overline{\sigma}) + (1 - a)\overline{\sigma} \\
&= a\lambda\sigma_i + a(1 - \lambda)\overline{\sigma} + (1 - a)\overline{\sigma} \\
&= a\lambda\sigma_i + (1 - a\lambda)\overline{\sigma} = \hat{\sigma}_i(a\lambda).
\end{aligned}$$

Proof of Proposition 6

Consider a new density that places probability $p(\sigma_i)$ on σ_i where:

$$p(\sigma_1) = \sum_{i=1}^{n} f(\sigma_i)\frac{\sigma_n - \sigma_i}{\sigma_n - \sigma_1}$$
$$p(\sigma_i) = 0 \quad i \in \{2,\ldots,n-1\}.$$
$$p(\sigma_n) = \sum_{i=1}^{n} f(\sigma_i)\frac{\sigma_i - \sigma_1}{\sigma_n - \sigma_1}$$

Note that $\Sigma_{i=1}^{n} p(\sigma_i) = \Sigma_{i=1}^{n} f(\sigma_i)((\sigma_n - \sigma_i)/(\sigma_n - \sigma_1) + (\sigma_i - \sigma_1)/(\sigma_n - \sigma_1)) = \Sigma_{i=1}^{n} f(\sigma_i) = 1$. Moreover, the mean is preserved as:

$$\begin{aligned}\sum_{i=1}^{n} p(\sigma_i)\sigma_i &= p(\sigma_1)\sigma_1 + p(\sigma_n)\sigma_n\\ &= \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_n - \sigma_1}(\sigma_1(\sigma_n - \sigma_i) + \sigma_n(\sigma_i - \sigma_1))\\ &= \sum_{i=1}^{n} f(\sigma_i)\sigma_i = \bar{\sigma}.\end{aligned}$$

We will now show that the new distribution is a mean-preserving spread of the original one. Let $P(\sigma_i) = \Sigma_{j=1}^{i} p(\sigma_j)$. Recall that when two distributions share the same mean, P is a mean-preserving spread of F if and only if F second-order stochastically dominates P. Moreover, if P and F cross just once, with $P(\sigma_i) \geq F(\sigma_i)$ for all i less than some threshold i^* and $P(\sigma_i) \leq F(\sigma_i)$ for $i > i^*$, then F second-order stochastically dominates P (see Fishburn and Vickson 1978, section 2.16). Hence, it is sufficient to establish that F crosses P once from below to prove that P is a mean-preserving spread of F. To this end, we have $P(\sigma_1) = p(\sigma_1) = f(\sigma_1) + \Sigma_{i=2}^{n} f(\sigma_i)(\sigma_n - \sigma_i)/(\sigma_n - \sigma_1) > f(\sigma_1) = F(\sigma_1)$. Because $P(\sigma_i)$ is flat for $i < n$, whereas F is increasing, to establish that F and P cross once before σ_n, it is sufficient to demonstrate that $P(\sigma_{n-1}) - F(\sigma_{n-1}) < 0$. We have:

$$\begin{aligned}P(\sigma_{n-1}) - F(\sigma_{n-1}) &= \sum_{i=1}^{n-1} f(\sigma_i)\left(\frac{\sigma_n - \sigma_i}{\sigma_n - \sigma_1} - 1\right)\\ &= \sum_{i=1}^{n-1} f(\sigma_i)\left(\frac{\sigma_1 - \sigma_i}{\sigma_n - \sigma_1}\right) < 0.\end{aligned}$$

Thus, P is a mean-preserving spread of F. To complete the proof, we must show that under the density $p(\cdot)$, the owner prefers an informed agent to an uninformed agent when $2\sigma_1 \leq \sigma_n$. From Proposition 4, this requires (letting $p_i = p(\sigma_i)$):

$$\frac{p_1}{\sigma_1} + \frac{1 - p_1}{\frac{1}{1 - p_1}(\sigma_n - \sigma_1) + \sigma_1} > \frac{1}{\bar{\sigma}},$$

which is equivalent to any of the following statements:

$$\frac{p_1}{\sigma_1} + \frac{(1-p_1)^2}{\sigma_n - \sigma_1 + \sigma_1(1-p_1)} > \frac{1}{p_1\sigma_1 + (1-p_1)\sigma_n}$$

$$\frac{p_1(\sigma_n - p_1\sigma_1) + (1-p_1)^2\sigma_1}{\sigma_1(\sigma_n - \sigma_1 p_1)} > \frac{1}{p_1\sigma_1 + (1-p_1)\sigma_n}$$

$$p_1^2\sigma_1\sigma_n - p_1^3\sigma_1^2 + (1-p_1)^2 p_1\sigma_1^2 + p_1(1-p_1)\sigma_n(\sigma_n - p_1\sigma_1) + (1-p_1)^3\sigma_1\sigma_n$$
$$> \sigma_1\sigma_n - p_1\sigma_1^2$$

$$(1-p_1)^2\sigma_1^2 p_1 + p_1(1-p_1)\sigma_n(\sigma_n - p_1\sigma_1) + (1-p_1)^3\sigma_1\sigma_n$$
$$> (\sigma_1\sigma_n - p_1\sigma_1^2)(1-p_1)(1+p_1)$$

$$(1-p_1)\sigma_1^2 p_1 + p_1\sigma_n(\sigma_n - p_1\sigma_1) + (1-p_1)^2\sigma_1\sigma_n > (\sigma_1\sigma_n - p_1\sigma_1^2)(1+p_1)$$

$$(1-p_1)\sigma_1^2 p_1 + p_1\sigma_n(\sigma_n - p_1\sigma_1) + p_1\sigma_1^2(1+p_1) + \sigma_1\sigma_n(p_1(p_1-3)) > 0$$

$$(1-p_1)\sigma_1^2 + \sigma_n(\sigma_n - p_1\sigma_1) + \sigma_1^2(1+p_1) + \sigma_1\sigma_n(p_1-3) > 0$$

$$\sigma_1^2(1 - p_1 + k(k-p_1) + 1 + p_1 + k(p_1-3)) > 0 \quad \left(k = \frac{\sigma_n}{\sigma_1}\right)$$

$$(k-2)(k-1) > 0.$$

Proof of Proposition 7

The owner solves:

$$\max_{\mathbf{x},\mathbf{y}} \sum_{i=1}^{n} f(\sigma_l)((1-y_l)r_{l,l} - x_l) \tag{P2}$$

such that:

$$y_l r_{l,l} + x_l - c_i(r_{l,l}) \geq 0 \quad \forall i \tag{17}$$

$$y_i r_{i,i} + x_i - c_i(r_{i,i}) \geq y_l r_{i,l} + x_l - c_i(r_{i,l}) \quad \forall i,l. \tag{18}$$

As before (17) will bind only for the worst type ($i = n$). We will solve the problem assuming that the local upward constraints in (18) bind, and then show that the optimal solution satisfies all the IC constraints in (18).

As discussed in the text, a type i manager facing contract $y_l r + x_l$ will choose to generate an expected revenue of $r_{i,l} = (\boldsymbol{\beta} \cdot \boldsymbol{\beta})\, y_l/\hat{\sigma}_i$. Since (18) binds for type n, set:

$$x_n = -\frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{2\hat{\sigma}_n} y_n^2.$$

As the local upward ICs bind for $i < n$, we can write:

$$x_l = y_{l+1}r_{l,l+1} + x_{l+1} - c_l(r_{l,l+1}) - y_l r_{l,l} + c_l(r_{l,l})$$
$$= x_{l+1} + \frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})}{2\hat{\sigma}_l}(y_{l+1}^2 - y_l^2).$$

Substituting back into the owner's objection function, we can simplify P2 to:

$$\max_{\mathbf{y}}(\boldsymbol{\beta} \cdot \boldsymbol{\beta})\left(\sum_{i=1}^{n} f(\sigma_i)\,\frac{(1-y_i)y_i}{\hat{\sigma}_i} + \sum_{i=1}^{n} F(\sigma_i)\,\frac{y_i^2}{2\hat{\sigma}_i} - \sum_{i=2}^{n} F(\sigma_{i-1})\,\frac{y_i^2}{2\hat{\sigma}_{i-1}}\right). \tag{19}$$

Solving, we find that the owner optimally sets:

$$y_i = \frac{1}{\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1}.$$

We claim that y_i is decreasing in i. Assumption A1 implies that $F(\sigma_i)/f(\sigma_i)$ is increasing in i. Therefore, it is sufficient to show that $\hat{\sigma}_i/\hat{\sigma}_{i-1}$ is nondecreasing in i. By A2*, σ_i/σ_{i-1} is increasing in i. If $a = 0$, then $\hat{\sigma}_i = \sigma_i$ and we are done. If $a > 0$, then we can write:

$$\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} = \frac{a\sigma_i + (1-a)\overline{\sigma}}{a\sigma_{i-1} + (1-a)\overline{\sigma}} = \frac{\sigma_i + S}{\sigma_{i-1} + S}.$$

Now, $\hat{\sigma}_i/\hat{\sigma}_{i-1}$ is increasing in i if and only if any of the following equivalent statements are true:

$$\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} \geq \frac{\hat{\sigma}_{i-1}}{\hat{\sigma}_{i-2}}$$

$$\frac{\sigma_i + S}{\sigma_{i-1} + S} \geq \frac{\sigma_{i-1} + S}{\sigma_{i-2} + S}$$

$$(\sigma_i + S)(\sigma_{i-2} + S) - (\sigma_{i-1} + S)^2 \geq 0. \tag{20}$$

Now, (20) holds trivially for $S = 0$ by A2*. To see that (20) holds for $S > 0$, we can differentiate (20) in S to obtain:

$$\frac{\partial}{\partial S}\left((\sigma_i + S)(\sigma_{i-2} + S) - (\sigma_{i-1} + S)^2\right) = \sigma_i + \sigma_{i-2} - 2\sigma_{i-1} > 0.$$

The inequality above holds by A2* and the following claim:

Claim: Suppose $a > b > c > 0$ and $a/b \geq b/c$, then $a + c - 2b > 0$.

To see this, note that $a/b \geq b/c \Rightarrow ac > b^2$. Therefore $a + c - 2b > a + c - 2\sqrt{ac} = (\sqrt{a} - \sqrt{c})^2 > 0$.

We now verify that the optimal contract satisfies global upward IC via the following equivalent inequalities:

$$x_i + y_i r_{i,i} - c_i(r_{i,i}) \geq x_{i+l} + y_{i+l} r_{i,i+l} - c_i(r_{i,i+l})$$

$$x_i \geq x_{i+l} + \frac{\boldsymbol{\beta} \cdot \boldsymbol{\beta}}{2\hat{\sigma}_i}(y_{i+l}^2 - y_i^2)$$

$$\sum_{t=i}^{i+l-1} \frac{1}{2\hat{\sigma}_{t+1}}(y_{t+1}^2 - y_t^2) \geq \frac{1}{2\hat{\sigma}_i}(y_{i-l}^2 - y_i^2).$$

The last inequality above follows from the monotonicity of y_i and $\hat{\sigma}_i$ in i. To verify downward IC, note that the following inequalities are equivalent:

$$x_i + y_i r_{i,i} - c_i(r_{i,i}) \geq x_{i-l} + y_{i-l} r_{i,i-l} - c_i(r_{i,i-l})$$

$$x_i \geq x_{i-l} + \frac{\boldsymbol{\beta} \cdot \boldsymbol{\beta}}{2\hat{\sigma}_i} (y_{i-l}^2 - y_i^2)$$

$$0 \geq \sum_{t=i-l}^{i-1} \frac{1}{2\hat{\sigma}_{t+1}} (y_{t+1}^2 - y_t^2) + \frac{\boldsymbol{\beta} \cdot \boldsymbol{\beta}}{2\hat{\sigma}_i} (y_{i-l}^2 - y_i^2).$$

The last inequality, again, holds by the monotonicity of $\hat{\sigma}_i$ and y_i in i.

Proof of Proposition 8

Recall from Proposition 7 that the owner's unconstrained objective function is:

$$M(\mathbf{y}, a) = (\boldsymbol{\beta} \cdot \boldsymbol{\beta})\left(\sum_{i=1}^{n} (\sigma_i - \overline{\sigma})\left(f(\sigma_i)\,\frac{(1 - y_i)y_i}{\hat{\sigma}_i} + F(\sigma_i)\,\frac{y_i^2}{2\hat{\sigma}_i}\right)\right.$$
$$\left. - \sum_{i=2}^{n} (\sigma_{i-1} - \overline{\sigma})\, F(\sigma_{i-1})\,\frac{y_i^2}{2\hat{\sigma}_{i-1}}\right).$$

Let $V(a) = \max_{\mathbf{y}} M(\mathbf{y}, a)$. By the envelope theorem, $V'(0) = M_a(\mathbf{y}, 0)$.

$$M_a(\mathbf{y}, 0) = -(\boldsymbol{\beta} \cdot \boldsymbol{\beta})\left(\sum_{i=1}^{n} (\sigma_i - \overline{\sigma})\left(f(\sigma_i)\,\frac{(1 - y_i)y_i}{\hat{\sigma}_i^2} + F(\sigma_i)\,\frac{y_i^2}{2\hat{\sigma}_i^2}\right)\right.$$
$$\left. - \sum_{i=2}^{n} (\sigma_{i-1} - \overline{\sigma})F(\sigma_{i-1})\,\frac{y_i^2}{2\hat{\sigma}_{i-1}^2}\right).$$

If $a = 0$, then $\hat{\sigma}_i = \overline{\sigma}$ and $y_i = y$. Letting $F(\sigma_0) = 0$, we can write:

$$\begin{aligned} M_a(\mathbf{y}, 0) &= -\frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})y^2}{2\overline{\sigma}^2}\left(\sum_{i=1}^{n} (\sigma_i - \overline{\sigma})F(\sigma_i) - \sum_{i=2}^{n} (\sigma_{i-1} - \overline{\sigma})F(\sigma_{i-1})\right) \\ &= -\frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})y^2}{2\overline{\sigma}^2}\left(\sum_{i=1}^{n} (\sigma_i - \overline{\sigma})(F(\sigma_{i-1}) + f(\sigma_i)) - \sum_{i=2}^{n} (\sigma_{i-1} - \overline{\sigma})F(\sigma_{i-1})\right) \\ &= -\frac{(\boldsymbol{\beta} \cdot \boldsymbol{\beta})y^2}{2\overline{\sigma}^2} \sum_{i=2}^{n} F(\sigma_{i-1})(\sigma_i - \sigma_{i-1}) < 0, \end{aligned}$$

where the last inequality holds by the monotonicity of σ_i in i. To see that $V(a)$ is convex, we have:

$$\begin{aligned} M_{aa}(\mathbf{y},a) &= (\boldsymbol{\beta} \cdot \boldsymbol{\beta})\left(\sum_{i=1}^{n} \tau_i\left(f(\sigma_i)\,\frac{(1 - y_i)y_i}{\hat{\sigma}_i^3} + F(\sigma_i)\,\frac{y_i^2}{2\hat{\sigma}_i^3}\right) - \sum_{i=1}^{n} \tau_{i-1}F(\sigma_{i-1})\,\frac{y_i^2}{2\hat{\sigma}_{i-1}^3}\right) \\ &= (\boldsymbol{\beta} \cdot \boldsymbol{\beta})\left(\sum_{i=1}^{n} \tau_i f(\sigma_i)\,\frac{(1 - y_i)y_i}{\hat{\sigma}_i^3} + \sum_{i=1}^{n} \tau_{i-1}F(\sigma_{i-1})\left(\frac{y_{i-1}^2 - y_i^2}{2\hat{\sigma}_{i-1}^3}\right) + \frac{y_n^2\tau_n}{2\hat{\sigma}_n^3}\right), \end{aligned}$$

where $\tau_i = (\sigma_i(a) - \overline{\sigma})^2$. All the terms in the expression above are positive by the monotonicity of y_i, hence $V''(a) > 0$ by the claim in our proof to Proposition 3. The owner chooses $a = 0$ or $a = 1$, because she maximizes $V(a)$, which is convex in a.

Proof of Proposition 9

From Proposition 7, the principal's profits when the manager is uninformed is given by $\pi^I(0) = \frac{\beta \cdot \beta}{2\bar{\sigma}}$. Therefore, the principal prefers the agent to be perfectly informed as long as any of the following equivalent statements hold:

$$\begin{aligned}
\pi^I(0) &\leq \pi^I(1) \\
\frac{2}{\beta \cdot \beta}\pi^I(0) &\leq \frac{2}{\beta \cdot \beta}\pi^I(1) \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} 2y_i(1 - y_i)\frac{f(\sigma_i)}{\sigma_i} + \sum_{i=1}^{n} F(\sigma_i)\frac{y_i^2}{\sigma_i} - \sum_{i=1}^{n} F(\sigma_{i-1})\frac{y_i^2}{\sigma_{i-1}} \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}\left[2y_i(1 - y_i) + \frac{F(\sigma_i)}{f(\sigma_i)}y_i^2 - \frac{F(\sigma_{i-1})}{f(\sigma_i)}y_i^2\frac{\sigma_i}{\sigma_{i-1}}\right] \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}\left[2y_i - 2y_i^2 + \frac{F(\sigma_i)}{f(\sigma_i)}y_i^2 - \frac{F(\sigma_i)}{f(\sigma_i)}y_i^2\frac{\sigma_i}{\sigma_{i-1}} + y_i^2\frac{\sigma_i}{\sigma_{i-1}}\right] \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}\left[2y_i - y_i^2 - y_i^2\left(1 - \frac{\sigma_i}{\sigma_{i-1}}\right) + \frac{F(\sigma_i)}{f(\sigma_i)}y_i^2\left(1 - \frac{\sigma_i}{\sigma_{i-1}}\right)\right] \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}\left[2y_i - y_i^2 - y_i^2\left(1 - \frac{\sigma_i}{\sigma_{i-1}}\right)\left(1 - \frac{F(\sigma_i)}{f(\sigma_i)}\right)\right].
\end{aligned}$$

For $a = 1$, the closed-form solution given for y_i in Proposition 7 implies that $\frac{1 - y_i}{y_i} = (1 - (\sigma_i/\sigma^{i-1}))(1 - (F(\sigma_i)/f(\sigma_i)))$. Therefore, we can rewrite the last line above as:

$$\begin{aligned}
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}\left[2y_i - y_i^2 - y_i^2\frac{1 - y_i}{y_i}\right] \\
\frac{1}{\bar{\sigma}} &\leq \sum_{i=1}^{n} \frac{f(\sigma_i)}{\sigma_i}[y_i] \\
\frac{1}{\bar{\sigma}} &\leq \mathrm{E}\left[\frac{y_i}{\sigma_i}\right].
\end{aligned}$$

Proof of Proposition 10

From Proposition 7, we have:

$$y_i'(a) = \frac{\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)(\hat{\sigma}_{i-1}(a)\hat{\sigma}_i'(a) - \hat{\sigma}_{i-1}'(a)\hat{\sigma}_i(a))}{\hat{\sigma}_{i-1}^2(a)\left(\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1\right)^2}.$$

Note that we can write:

$$\begin{aligned}\hat{\sigma}_{i-1}(a)\hat{\sigma}_i'(a) - \hat{\sigma}_{i-1}'(a)\hat{\sigma}_i(a) &= (a\sigma_{i-1} + (1 - a)\bar{\sigma})(\sigma_i - \bar{\sigma}) \\ &\quad - (\sigma_{i-1} - \bar{\sigma})(a\sigma_i + (1 - a)\bar{\sigma}) \\ &= (\sigma_i - \sigma_{i-1})\bar{\sigma}.\end{aligned}$$

As such, we can re-express $y_i'(a)$ as:

$$y_i'(a) = -\frac{\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\bar{\sigma}(\sigma_i - \sigma_{i-1})}{\hat{\sigma}_{i-1}^2(a)\left(\left(\frac{F(\sigma_i)}{f(\sigma_i)} - 1\right)\left(\frac{\hat{\sigma}_i}{\hat{\sigma}_{i-1}} - 1\right) + 1\right)^2} \leq 0,$$

where strict inequality holds for all $i > 1$ since $\sigma_i > \sigma_{i-1}$ and $F(\sigma_i) > f(\sigma_i)$.

Proof of Lemma 2

Note that:

$$\begin{aligned}W(a) = \nu\frac{1}{2\sigma_1}\Bigg(&\frac{1}{\rho(1 - a) + 1 + a} + \frac{2}{3a(\rho - 1) + \rho + 1} \\ &- \frac{a(\rho - 1) + \rho + 1}{(3a(\rho - 1) + \rho + 1)^2}\Bigg).\end{aligned}$$

Therefore:

$$W'(a) = \nu\frac{1}{2\sigma_1}\left(\frac{4a(\rho - 1)(3a^2(\rho - 1)^2 - (\rho + 1)^2 + 14a(\rho^2 - 1))}{(1 + 3a(\rho - 1) + \rho)^3(1 + a + \rho - a\rho)^2}\right).$$

The equation $W'(a) = 0$ has three solutions: $a = 0$ and the two solutions to the quadratic, $3a^2(\rho - 1)^2 + 14a(\rho^2 - 1) - (\rho + 1)^2 = 0$. Consider first $a = 0$. We have:

$$W''(0) = -\frac{2\nu(\rho - 1)^2}{(\rho + 1)^3\sigma_1} < 0.$$

Hence $W(a)$ must be initially decreasing at $a = 0$. The additional two roots are $\frac{\rho + 1}{3(\rho - 1)}(-7 \pm 2\sqrt{13})$, one of which is clearly negative. Label $a^+ = \frac{\rho + 1}{3(\rho - 1)}(-7 + 2\sqrt{13})$. If $a^+ \geq 1$, then the function $W(a)$ is decreasing over the entire interval $[0,1]$; if $a^+ < 1$, then the function $W(a)$ is decreasing on the range $[0,a^+]$ and increases over $[a^+,1]$. In either case, it is clear that the function is maximized at either $a = 0$ or $a = 1$. Plugging in these values into the closed-form solution for $W(a)$, we find that $W(1) \geq W(0)$ if and only if:

$$\frac{(\rho - 1)^2(4\rho - 5)}{8\rho^3 - 6\rho + 2} \geq 0,$$

which is equivalent to $\rho \geq 1.25$.

Proof of Proposition 11

We will show that $R(a)$ is single peaked on the interval $[0,2(1 + \rho)/3(\rho - 1)]$. Note that $a = \dfrac{1 + \rho}{3(\rho - 1)}$ is the unique solution to $R'(a) = 0$. Moreover, we have:

$$R''(a) = \frac{6(\rho - 1)(3a(\rho - 1) - 2(\rho + 1))}{(1 + \rho + 3a(\rho - 1))^4},$$

which is negative in the range $a \in \left[0, \dfrac{2(\rho + 1)}{3(\rho - 1)}\right]$. As such, the manager will select $a = 1$ whenever $\dfrac{\rho + 1}{3(\rho - 1)} \geq 1$, i.e., when $\rho \leq 2$. When $\rho > 2$, the manager's rents are maximized at $a = \dfrac{\rho + 1}{3(\rho - 1)} < 1$.

Proof of Lemma 3

Given uniform density, Proposition 9 implies that the owner prefers $a = 1$ over $a = 0$ if and only if:

$$\sum_{i=1}^{n} \frac{1}{\rho^{i-1}(1 + (i - 1)(\rho - 1))} \geq \frac{n^2(\rho - 1)}{\rho^2 - 1}.$$

For $n = 2$, this reduces to the following equivalent statements:

$$\rho^3 - 3\rho^2 + \rho + 1 \geq 0$$
$$\rho \geq 1 + \sqrt{2}.$$

Therefore, the principal prefers $a = 0$ when $\rho \in (1,1 + \sqrt{2}]$. Now, denote the manager's rents under the menu of linear contracts by $R_l(a)$ where:

$$R_l(a) = \frac{1}{2(\beta \cdot \beta)} \sum_{i=1}^{n} y_i(a)^2\left(\frac{1}{\hat{\sigma}_{i-1}} - \frac{1}{\hat{\sigma}_i}\right)F(\sigma_{i-1}).$$

When $n = 2$, the rents simplify to:

$$\begin{aligned}
R_l(a) &= \frac{1}{2(\beta \cdot \beta)} \frac{\hat{\sigma}_1}{\hat{\sigma}_2^3}\, a(\rho - 1)\sigma_1 \\
&= \frac{(\rho - 1)\sigma_1}{2(\beta \cdot \beta)} \frac{a(\sigma_1 + (1 - a)\overline{\sigma})}{(a\sigma_2 + (1 - a)\overline{\sigma})^3} \\
&= \frac{(\rho - 1)\sigma_1}{2(\beta \cdot \beta)} \frac{a\frac{\sigma_1}{2}(2a + 1 - a + \rho(1 - a))}{\left(\frac{\sigma_1}{2}\right)^3(2a\rho + (1 - a)(\rho + 1))^3} \\
&= \kappa \frac{a(a + 1 + \rho(1 - a))}{(\rho + 1 + a(\rho - 1))^3},
\end{aligned}$$

where κ is a positive constant. Differentiating, we have:

$$R_l'(a) = \kappa \frac{a^2(\rho - 1)^2 + (\rho + 1)^2 - 4a(\rho^2 - 1)}{(\rho + 1 + a(\rho - 1))^4}.$$

The numerator above is positive at $a = 0$ and decreases in a for $a \in [0,1]$. If $\rho \leq \sqrt{3}$, the numerator of $R_l'(1)$ is non-negative, implying that $R_l'(a) > 0$ for $a \in [0,1)$; as such, the manager would prefer $a = 1$. For $\rho > \sqrt{3}$, there is a unique interior solution at which the first-order condition is met. Solving, we find that the manager's optimal choice is given by $a = (2 - \sqrt{3})\,\frac{\rho + 1}{3(\rho - 1)} < 1$.

Overall welfare can be expressed as $W_l(a)$, where:

$$W_l(a) = \frac{1}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} \sum_{i=1}^{n} \frac{y_i(2 - y_i)}{\hat{\sigma}_i} f(\sigma_i).$$

With $n = 2$ and geometric states, we have:

$$W_l(a) = \frac{1}{2(\boldsymbol{\beta} \cdot \boldsymbol{\beta})} \left(\frac{1}{\hat{\sigma}_1} + \frac{\hat{\sigma}_1(2\hat{\sigma}_2 - \hat{\sigma}_1)}{\hat{\sigma}_2^3}\right).$$

Taking the derivative yields:

$$W_l'(a) =$$
$$C\,\frac{2a(\rho - 1)^2[a^3(\rho - 1)^3 - 3a^2(\rho - 1)^2(\rho + 1) + 7a(\rho - 1)(\rho + 1)^2 - (1 + \rho)^3]}{(1 + a + \rho(1 - a))^2(1 - a + \rho(1 + 4))^4}.$$

Note that $W_l'(0) = 0$, but $W_l''(0) = -2(\rho - 1)^2/(\rho + 1)^3 < 0$, so $W_l'(a)$ is decreasing at $a = 0$. Moreover, the cubic in the numerator of $W_l'(a)$ has a unique real root; hence, $W_l(a)$ changes direction at most once over $a \in [0,1]$, at a global minimum. Therefore, $W_l(a)$ must be maximized at either $a = 0$ or $a = 1$. Comparing $W_l(0)$ with $W_l(1)$ shows that $W_l(1) \geq W_l(0)$ if and only if $\rho \geq (1 + \sqrt{5})/2$.

REFERENCES

Antle, R., and J. Fellingham. 1995. Information rents and preferences among information systems in a model of resource allocation. *Journal of Accounting Research* 33 (Supplement): 41–58.

Arya, A., J. Glover, and K. Sivaramakrishnan. 1997. The interaction between decision and control problems and the value of information. *The Accounting Review* 72 (October): 561–574.

Bagnoli, M., and T. Bergstrom. 2005. Log-concave probability and its applications. *Economic Theory* 26: 445–469.

Baiman, S., and J. Evans III. 1983. Pre-decision information and participative management control systems. *Journal of Accounting Research* 21 (Autumn): 371–395.

———, and K. Sivaramakrishnan. 1991. The value of private pre-decision information in a principal-agent context. *The Accounting Review* 66 (October): 747–766.

Baker, G., and B. Jorgensen. 2003. Volatility, noise, and incentives. Working paper, Harvard Business School, Cambridge, MA.

Banker, R., and S. Datar. 1989. Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research* 27 (1): 21–39.

Baron, D., and R. Myerson. 1982. Regulating a monopolist with unknown costs. *Econometrica* 50 (July): 11–30.

Blackwell, D. 1951. *Comparison of experiments*. Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 93–102.

Bushman, R., R. Indjejikian, and M. Penno. 2000. Private predecision information, performance measure congruity, and the value of delegation. *Contemporary Accounting Research* 17 (Winter): 561–587.

Cachon, G., and M. Lariviere. 2001. Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science* 47 (May): 629–646.

Christensen, J. 1981. Communication in agencies. *Bell Journal of Economics* 12 (2): 661–674.

Cremer, J., and F. Khalil. 1992. Gathering information before signing a contract. *American Economic Review* 82 (June): 566–578.

———, ———, and J-C. Rochet. 1998. Strategic information gathering before a contract is offered. *Journal of Economic Theory* 81 (1): 163–200.

Datar, S., S. Kulp, and R. Lambert. 2001. Balancing performance measures. *Journal of Accounting Research* 39 (June): 75–92.

Fishburn, P., and R. Vickson. 1978. Theoretical foundations of stochastic dominance. In *Stochastic Dominance*, edited by M. C. Findlay, and Whitmore, 39–113. Lanham, MD: Lexington Books.

Holmstrom, B. 1979. Moral hazard and observability. *Bell Journal of Economics* 10 (Spring): 74–91.

Kessler, A. 1998. The value of ignorance. *Rand Journal of Economics* 29 (Summer): 339–354.

Kirby, A., S. Reichelstein, P. Sen, and T-Y. Paik. 1991. Participation, slack, and budget-based performance evaluation. *Journal of Accounting Research* 29 (Spring): 109–128.

Lewis, T., and D. Sappington. 1991. All-or-nothing information control. *Economics Letters* 37 (October): 111–113.

———, and ———. 1993. Ignorance in agency problems. *Journal of Economic Theory* 61 (1): 169–183.

Maskin, E., and J. Riley. 1984. Monopoly with incomplete information. *RAND Journal of Economics* 15 (Summer): 171–196.

Penno, M.. 1984. Asymmetry of pre-decision information and managerial accounting. *Journal of Accounting Research* 22 (Spring): 171–191.

Raith, M. 2004. *Specific Knowledge and Performance Measurement*. CEPR Discussion Paper No. 4262. London, U.K.: Centre for Economic Policy Research.

Reichelstein, S. 1992. Constructing incentive schemes for government contracts: An application of agency theory. *The Accounting Review* 67 (October): 712–731.

Rogerson, W. P. 1987. On the optimality of menus of linear contracts. Technical report, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, IL.

Sappington, D. 1984. Incentive contracting with asymmetric and imperfect precontractual knowledge. *Journal of Economic Theory* 34 (1): 52–70.

THE ACCOUNTING REVIEW
Vol. 81, No. 3
2006
pp. 713–743

The Implications of Accounting Distortions and Growth for Accruals and Profitability

Scott A. Richardson
University of Pennsylvania

Richard G. Sloan
University of Michigan

Mark T. Soliman
Stanford University

İrem Tuna
University of Pennsylvania

ABSTRACT: Following Sloan (1996), numerous studies document that the accrual component of earnings is less persistent than the cash flow component of earnings. Disagreement exists, however, as to the explanation for this result. One stream of literature follows Sloan's lead in arguing that this result is attributable to accounting distortions (Xie 2001; Dechow and Dichev 2002; Richardson et al. 2005). A second stream of literature argues that this result is attributable to a more general growth effect and that growth-related factors such as diminishing returns to new investment explain the lower persistence of accruals (e.g., Fairfield et al. 2003a; Cooper et al. 2005). We provide new evidence indicating that temporary accounting distortions are a significant contributing factor to the lower persistence of the accrual component of earnings. Our evidence indicates that the lower persistence of accruals extends to accruals that are unrelated to sales growth and that extreme accruals are systematically associated with alleged cases of earnings manipulation.

Keywords: *accruals; earnings management; conservative accounting; aggressive accounting.*

Data Availability: *The data used in this study are available from the sources identified in the study.*

We thank Patricia Dechow, Patricia Fairfield, Charles Lee, Mark Nelson, Terry Shevlin (editor), Teri Yohn, anonymous referees, and seminar participants at Cornell University, Georgetown University, University of Western Australia and participants at the 2004 AFAANZ conference for comments.

Editor's note: This paper was accepted by Terry Shevlin.

Submitted March 2004
Accepted November 2005

I. INTRODUCTION

Understanding the properties of accruals is arguably one of the most important goals of financial accounting research. The primary objective of financial accounting is to provide information that is useful to investors. Accrual accounting has emerged as the accepted method of achieving this objective. Accrual accounting centers on the identification and measurement of assets and liabilities, with accruals representing changes in noncash assets and liabilities. A natural goal for financial accounting research is to investigate how accruals facilitate investor decision making. In an attempt to address this question, Dechow (1994) provides evidence that accrual accounting earnings are superior to cash accounting earnings at summarizing firm performance. In follow-up research, Sloan (1996) shows that the accrual component of earnings is less persistent than the cash component of earnings. Sloan's (1996) results suggest that while accrual accounting is superior to cash accounting, the accrual component of earnings should receive a lower weighting than the cash component of earnings in evaluating firm performance. He attributes this result to the greater subjectivity involved in the estimation of accruals. Sloan (1996) also finds that investors do not appear to assign a lower weighting to accruals, leading to significant security mispricing. He concludes that investors do not fully comprehend the greater subjectivity involved in the estimation of accruals, causing them to make flawed investment decisions.

Subsequent research extends Sloan's (1996) work on the lower persistence of the accrual component of earnings. This research can be classified into two broad streams. The first stream of research builds on Sloan's (1996) conjecture that the lower persistence of the accrual component of earnings arises because the estimation of accruals involves greater subjectivity. For example, Xie (2001) shows that the lower persistence of the accrual component of earnings is attributable to the "abnormal" component of accruals, defined using the modified Jones (1991) model of normal accruals. He concludes that the lower persistence of accruals is primarily attributable to opportunistic managerial discretion. Dechow and Dichev (2002) and Richardson et al. (2005) (hereafter, RSST) show that the lower persistence of accruals is consistent with the existence of estimation error in accruals. These papers conclude that the lower persistence of the accrual component of earnings is attributable to temporary accounting distortions arising from accrual estimation error. Note that such distortions could arise both from unintentional errors in forecasting future benefits and obligations and from intentional managerial manipulation.

The second stream of research argues that accruals are correlated with economic characteristics, such as firm growth, and that these correlated economic characteristics are responsible for the lower persistence of the accrual component of earnings. In particular, Fairfield et al. (2003a) (hereafter, FWY) show that the lower persistence of the accrual component of earnings extends from the working capital accruals considered by Sloan (1996) to growth in noncurrent balance sheet accounts. They suggest that the lower persistence of the accrual component of earnings arises from the interaction of firm growth with the lower rates of economic profits associated with diminishing marginal returns to increased investment. Variants of this explanation are embraced by Titman et al. (2004), Anderson and Garcia-Feijoo (2006), and Cooper et al. (2005). FWY also propose that the lower persistence of the accrual component of earnings may result from the interaction of firm growth with the conservative bias in generally acceptable accounting principles.

The objective of this paper is to provide additional evidence on the role of accounting distortions in explaining the lower persistence of the accrual component of earnings. We begin with a series of numerical examples that illustrate the impact of accounting distortions on accruals and earnings. The key insight emerging from these examples is that the observed empirical relation between accruals and earnings persistence is consistent with the existence of temporary accounting distortions arising from accrual estimation error. Next, we provide two sets of empirical tests examining the role of temporary accounting distortions. First, we decompose accruals into a "growth" component and an "efficiency" component. The growth component reflects accruals that are attributable to growth in output, while the efficiency component reflects accruals that are unrelated to growth in output. The effects of diminishing marginal returns to increased investment should be captured by the growth component of accruals. Our results indicate that both the growth and efficiency components contribute to the lower persistence of earnings. These results indicate that diminishing marginal returns do not provide a complete explanation for the lower persistence of the accruals, suggesting that accounting distortions play a significant role.

Our second set of tests examines the association between accruals and SEC enforcement actions for alleged earnings manipulations. We show that firms subject to SEC enforcement actions have abnormally high accruals at the time of the alleged earnings manipulations and unusually low accruals following the alleged earnings manipulations. These results are consistent with the opportunistic use of managerial discretion to generate accounting distortions that temporarily inflate accruals and earnings.

Overall, our analysis and results suggest that temporary accounting distortions provide the most compelling explanation for the lower persistence of the accrual component of earnings. Moreover, our SEC enforcement action tests suggest that some of these distortions arise from the opportunistic use of managerial discretion. At the same time, we cannot rule out a supplementary role for growth-related explanations, such as diminishing marginal returns to new investment.

Our findings have important implications for financial accounting, because they indicate that the lower persistence of accruals is at least partially attributable to distortions arising from accrual accounting. Accrual accounting involves a trade-off between the increased relevance associated with accruing future benefits and obligations and the reduced reliability associated with estimating the amounts of these future benefits and obligations. Our results suggest that the reduced reliability associated with accrual accounting results in temporary accounting distortions that have a significant impact on earnings persistence. The current push by accounting standard setters toward "fair value" accounting is likely to increase the frequency and magnitude of accounting distortions, leading to a reduction in the persistence and reliability of earnings.

The remainder of the paper is organized as follows. Section II provides a more detailed discussion of accruals and their implications for earnings persistence. Section III develops numerical examples that illustrate the effects of accounting distortions on the properties of accruals and earnings. Section IV describes our data, Section V presents our results, and Section VI concludes.

II. RESEARCH DESIGN

Alternative Explanations for the Lower Persistence of the Accrual Component of Earnings

Our objective in this paper is to discriminate between alternative explanations for the lower persistence of the accrual component of accounting earnings (where earnings are deflated by an appropriate measure of invested capital and thus expressed as an accounting rate of return). Formally, this result states that $\gamma_2 < 0$ in the following equation:

$$RNOA_{t+1} = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 ACC_t + \upsilon_{t+1} \tag{1}$$

where:

$RNOA_{t+1}$ = accounting rate of return for period $t+1$;
$RNOA_t$ = accounting rate of return for period t;
ACC_t = accrual component of the accounting rate of return for period t; and
υ_{t+1} = period $t+1$ innovation in $RNOA_{t+1}$ (assumed to be mean zero and uncorrelated with $RNOA_t$ and ACC_t).

From an intuitive perspective, this result manifests itself in the form of the following empirical regularity:

> **Empirical Regularity:** Firms with high (low) accruals in the current period tend to have a high (low) accounting rate of return in the current period, but tend to experience significant reductions (increases) in their accounting rate of return in the subsequent period.

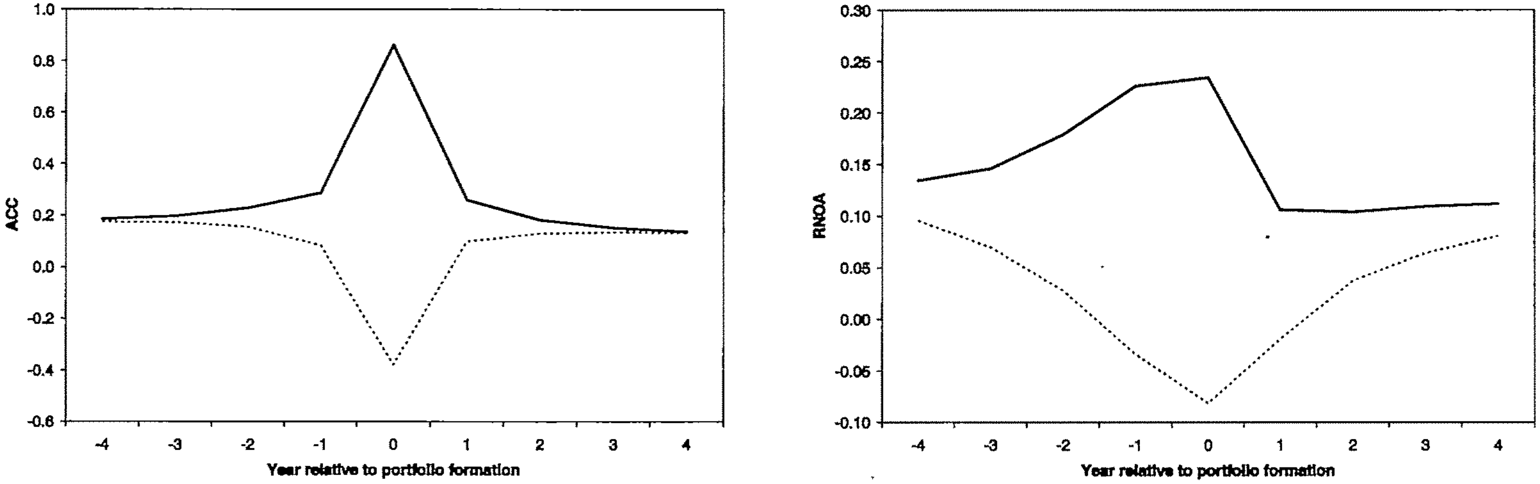
Panels A and B of Figure 1 illustrate this empirical regularity using our sample and variable definitions (we detail our sample formation and variable definitions later in the paper). This figure reports event-time plots for firm-years belonging to extreme accrual deciles in event-year 0. Panel A of Figure 1 plots the mean value of the accrual component of earnings for each of the extreme accrual deciles. This plot illustrates that the accrual component of earnings is rapidly mean reverting. Panel B of Figure 1 plots the mean value of the accounting rate of return for the extreme accrual deciles. This plot illustrates that firms in the high (low) accrual decile tend to have a high (low) accounting rate of return in event period 0, but tend to experience significant reductions (increases) in their accounting rate of return in the subsequent period. This is the regularity that we seek to explain in this paper. This regularity has proven to be robust across a variety of samples and variable definitions (see Sloan 1996; FWY; RSST).

Existing research offers two broad classes of explanations for the lower persistence of the accrual component of earnings. Following Sloan (1996), one line of research argues that it results from accounting distortions associated with the lower reliability of the accrual component of earnings. Following FWY, a second line of research argues that it is more likely to result from economic factors associated with growth in real investment and less likely to result from accounting distortions. The primary goal of our paper is to provide additional evidence that accounting distortions are a significant contributing factor to this empirical regularity.

Existing evidence relating to alternative explanations is indirect and inconclusive. Papers claiming to provide evidence in support of the accounting distortion explanation include Xie (2001), Dechow and Dichev (2002), and RSST. Xie decomposes accruals into "normal" and "abnormal" accruals using the modified Jones model (see Dechow et al. 1995). He finds that the lower persistence of earnings is primarily attributable to abnormal accruals. Based on this evidence, he concludes that the lower persistence of accruals is attributable to accounting distortions arising from earnings management. Dechow and Dichev (2002) attempt to isolate temporary accounting distortions by extracting the residuals from regressions of accruals on past, present, and future cash flows. They demonstrate that the lower persistence of the accrual component of earnings is concentrated in these residuals. RSST decompose accruals along broad balance sheet classifications and rank

FIGURE 1
Event-Time Plots of the Mean Values of Accruals (*ACC*), Accounting Rate of Return (*RNOA*), Sales Growth (*SG*), and Change in Asset Turnover (*ΔAT*) for Extreme Accrual (*ACC*) Deciles

Year 0 Represents the Accrual Ranking Year. The solid (dashed) line represents the highest (lowest) accrual decile.

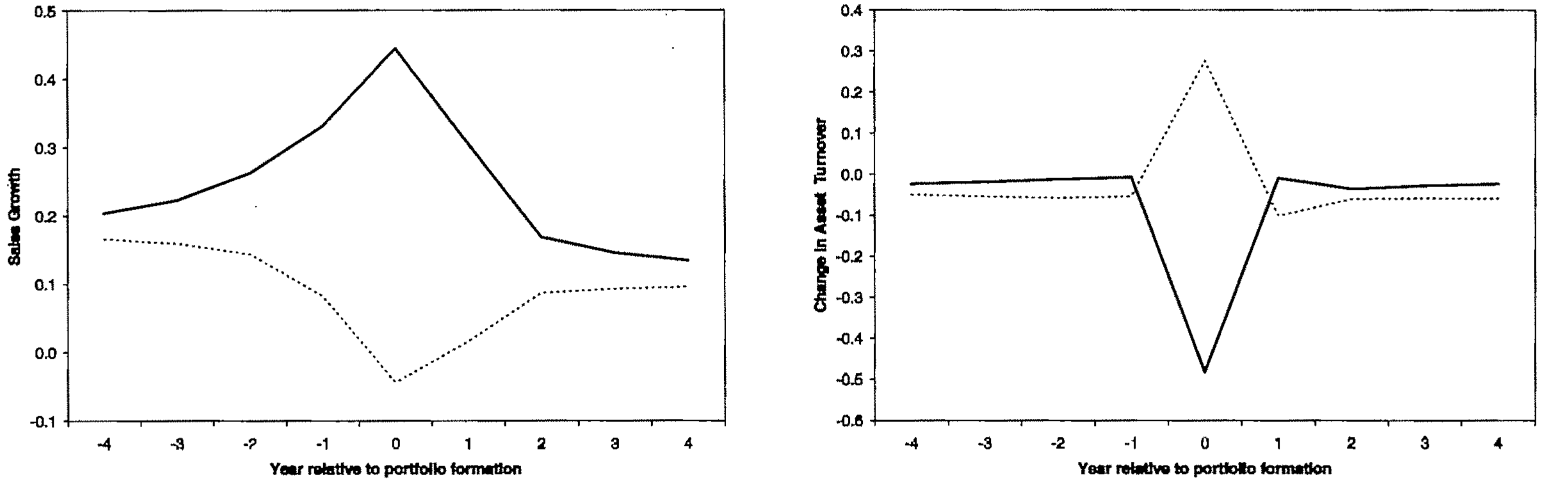


Panel A: Accruals (*ACC*)

Panel B: Accounting Rates of Return (*RNOA*)

(*continued on next page*)

FIGURE 1 (continued)



Panel C: Sales Growth (*SG*)

Panel D: Change in Asset Turnover (Δ*AT*) these classifications based on their assessment of the reliability of the underlying accrual categories. They show that the lower persistence of the accrual component of earnings is primarily attributable to accrual categories that they deem to be less reliable.

Papers proposing economic explanations for the lower persistence of the accrual component of earnings include FWY, Cooper et al. (2005), and Khan (2005). FWY show that the lower persistence of the accrual component of earnings extends from the working capital accruals considered by Sloan (1996) to growth in noncurrent operating assets. They suggest that the lower persistence of the accrual component of earnings arises from the interaction of investment growth with the lower rates of return from diminishing marginal returns to increased investment. In follow up research, Cooper et al. (2005) show that FWY's results extend to the growth rate in total assets. Following FWY, they also conclude that diminishing marginal returns to new investment is the most likely explanation. Finally, recent research by Khan (2005) and Zach (2005) argues that accruals proxy for risk. Khan (2005) argues that firms with high accruals better enable investors to smooth intertemporal consumption patterns, while firms with low accruals expose investors to distress risk. Accordingly, firms with high (low) accruals earn lower (higher) equilibrium profits and stock returns.

FWY also propose another possible explanation that combines both accounting distortions and economic growth. They argue that the widespread use of conservative accounting methods combines with growth in real investment to cause the lower persistence of the accrual component of earnings. We refer to this type of accounting distortion as a permanent accounting distortion, because it results from the consistent application of biased accounting methods. An example of such an accounting method would be the requirement to immediately and consistently expense research and development costs. Such permanent accounting distortions are subtly different from the previously discussed temporary accounting distortions. Temporary accounting distortions arise from transitory estimation error in accruals. Accruals represent estimates of future benefits and obligations. As such, they are more prone to estimation error than actual cash receipts or disbursements. An example of accrual estimation error would be the accrual of the allowance for uncollectible accounts on credit sales. A range of possible amounts could be feasibly booked, but only one of these amounts will reflect subsequent uncollectibles. All other amounts will contain transitory estimation error that is reversed once actual collections are known. We refer to this type of error as a temporary accounting distortion, because it results from transitory estimation error rather than the consistent application of biased accounting methods.

In order to provide evidence on the role of accounting distortions, our research design employs three complementary approaches. First, to illustrate the effects of accounting distortions on accruals and profitability, we develop a series of numerical examples. These examples assume constant marginal returns to investment, and so help us to isolate the impact of accounting distortions on accruals and profitability. They also allow us to determine which combinations of accounting distortions and growth are consistent with empirical regularities in accruals and profitability. Second, to discriminate between accounting distortions and diminishing marginal returns to new investment, we decompose accruals into a component attributable to growth in real investment ("growth component") and a component attributable to changes in the efficiency with which existing investments are employed ("efficiency component"). The effects of diminishing marginal returns should be limited to the growth component. Thus, to the extent that the efficiency component drives the observed properties of accruals, we infer that accounting distortions explain these properties. Third, to investigate the possibility that temporary accounting distortions are generated by managerial manipulation, we examine the relation between accruals and SEC enforcement actions for alleged earnings manipulations. The SEC issues an enforcement action in cases where it believes management has intentionally manipulated earnings in a manner that violates GAAP. Hence, evidence of a relation between SEC enforcement actions and accruals is consistent with temporary accounting distortions resulting from managerial manipulation.

Defining Accruals

Following Healy (1985), most academic research defines accruals as the change in noncash working capital less depreciation expense (e.g., Sloan 1996; Xie 2001; Thomas and Zhang 2001; FWY). This definition, however, omits accruals relating to noncurrent operating assets/liabilities and financial assets/liabilities. RSST propose a broader definition of accruals that incorporates all differences between GAAP earnings and cash earnings. They argue that in the absence of accrual accounting, only cash and owner's equity would appear on the balance sheet, because all other accounts are manifestations of the accrual accounting process. Under their definition, changes in all noncash asset and liability accounts represent accruals of expected future benefits or obligations. Total accruals therefore include working capital accruals (related to changes in noncash current operating asset/liability accounts), noncurrent operating accruals (related to changes in noncurrent operating asset/liability accounts), and financing accruals (related to changes in noncash financial asset/liability accounts).

RSST's expanded definition of accruals contrasts sharply with the definition of accruals used by FWY. FWY follow the established convention of referring only to working capital accruals as "accruals," and label noncurrent operating accruals as a generic form of "growth." FWY show that the change in net operating assets, which combines working capital accruals and noncurrent operating accruals, best captures the low-persistence component of earnings. They offer this result as their primary evidence against accounting-based explanations. The key difference, however, between working capital accruals and noncurrent operating accruals is that the future benefits and obligations associated with the latter usually take longer to be realized. Both are manifestations of the accrual accounting process, involving subjective estimates of future benefits and obligations. Moreover, as illustrated by the accounting debacle at WorldCom, noncurrent accruals are susceptible to earnings manipulation.[1] Thus, FWY's evidence does not help discriminate between accounting and economic-based explanations, because it is consistent with both.

RSST argue that both working capital accruals and noncurrent operating accruals are of low reliability and, hence, particularly prone to temporary accounting distortions. Consistent with these arguments, they show that the working capital and noncurrent operating accrual components of earnings are less persistent than the cash flow and financing accrual components of earnings. Thus, RSST reinterpret FWY's growth in net operating asset variable as a more comprehensive measure of accruals and FWY's results as a manifestation of temporary accounting distortions in accruals. Following RSST, we incorporate noncurrent operating assets into our definition of accruals, and we refer to the sum of working capital accruals and noncurrent operating accruals as total operating accruals. Thus, our definition of accruals includes all operating accruals, but excludes the more reliable financing accruals.

---

[1] WorldCom overstated its net income by more than $10 billion between 2000 and 2001, primarily by capitalizing line-operating costs in property, plant, and equipment. Dechow and Schrand (2004) and Nelson et al. (2003) also document evidence of earnings management through the inappropriate capitalization of operating costs. Surveys of earnings restatements by Huron Consulting Group also indicate that inappropriate capitalization of costs has been a major reason for restatements in recent years.

For expositional ease, we simply refer to this construct as the accrual component of earnings (denoted *ACC*) throughout the remainder of the paper. Note that given this definition of accruals, the corresponding definition of cash flows (i.e., the difference between earnings and accruals) is the standard finance textbook definition of free cash flow.[2] Exact variable definitions in terms of Compustat items are provided in Section IV.

Growth versus Efficiency Decomposition

FWY argue that diminishing marginal returns to new investment contribute to the lower persistence of the accrual component of earnings. In support of this argument, they provide a reference to Stigler (1963, 54). In this reference, Stigler asserts that:

> There is no more important proposition in economic theory than that, under competition, the rate of return on investment tends toward equality in all industries. Entrepreneurs will seek to leave relatively unprofitable industries and enter relatively profitable industries, and with competition, there will be neither public nor private barriers to these movements.

Note that FWY measure growth at the firm level while Stigler's (1963) arguments apply at the industry level. FWY's argument therefore requires that firm-level growth has a significant influence on industry-level output. This would only be the case if either the firm produces a significant proportion of industry output, or growth is strongly correlated across firms in the same industry.

In order to investigate the diminishing returns to new investment explanation in more detail, we decompose accruals into a "growth" component and an "efficiency" component. Accruals can increase either because of real investment growth (whereby more operating assets lead to increased production and sales), or because of reduced efficiency (whereby more operating assets are required to generate the same level of production and sales). Diminishing marginal returns to new investment will be reflected in the growth component of accruals. The efficiency component of accruals, in contrast, will pick up either accounting distortions or less efficient use of existing capital (see also, Jansen and Yohn 2003). For example, assume that inventories increase without corresponding increases in sales or payables. This increase will manifest itself as a decrease in the efficiency component of accruals. The two most likely explanations for such an increase are that inventories are overstated (i.e., a temporary accounting distortion) or that the firm is managing physical inventories less efficiently.

Existing research recognizes that real investment growth necessitates increased operating assets and so will be associated with higher accruals (e.g., Jones 1991). Such accruals are often termed "nondiscretionary" or "normal" in the academic literature. Nondiscretionary (discretionary) accruals are traditionally modeled as the fitted values (residuals) from regressions of accruals on measures of growth in operating activities, such as sales growth.[3] Recent studies by Xie (2001) and Chan et al. (2006) find that the lower persistence of the accrual component of earnings is primarily attributable to the discretionary component of accruals. However, the statistically oriented decompositions used in these studies are subject to misspecification. First, they impose an *ad hoc* linear structure on a nonlinear relation. Second, they require the estimation of key parameters (e.g., the sensitivity of accruals to sales growth). We use a parsimonious algebraic decomposition of accruals that mitigates these problems.

---

[2] Recall that free cash flow equals earnings, plus noncash charges such as depreciation, less changes in working capital, less investment expenditures. These adjustments to earnings correspond to the current and noncurrent operating accruals that we incorporate in our definition in accruals.

[3] See Dechow et al. (1995) for a more detailed discussion of the efficacy of models used for isolating the discretionary and nondiscretionary components of accruals.

The starting point for our growth versus efficiency decomposition is our measure of total operating accruals (*ACC*), defined as the change in net operating assets deflated by lagged operating assets. Net operating assets (*NOA*) is defined as the difference between noncash operating assets and operating liabilities. This definition corresponds to FWY's "growth in net operating assets" variable and RSST's "total operating accrual' variable."[4]

$$ACC_t = (NOA_t - NOA_{t-1})/NOA_{t-1}. \tag{2}$$

We next define efficiency using the standard net operating asset turnover ratio (*AT*):

$$AT_t = Sales_t/NOA_t. \tag{3}$$

The following decomposition then follows with some simple algebra:[5]

$$\begin{aligned} ACC_t &= \Delta NOA_t/NOA_{t-1} \\ &= \underbrace{\Delta \text{Sales}_t/Sales_{t-1}}_{\substack{\text{Sales}\\ \text{Growth}}} - \underbrace{\Delta AT_t/AT_t}_{\substack{\text{Change in}\\ \text{Efficiency}}} - \underbrace{(\Delta Sales_t/Sales_{t-1})^*(\Delta AT_t/AT_t)}_{\text{Interaction}}. \end{aligned} \tag{4}$$

Note that this decomposition is an algebraic identity that does not require the estimation of any additional parameters. The decomposition highlights the two determinants of operating accruals. First, accruals are directly related to sales growth. If asset efficiency remains unchanged, then sales growth will lead to proportional increases in accruals. Second, accruals are negatively related to efficiency. In the absence of sales growth, reductions in asset efficiency lead to proportional increases in accruals. The decomposition also contains an interaction term, indicating that a simple linear decomposition is not appropriate when sales growth and efficiency changes are correlated. A nonzero correlation is a distinct possibility. Economies of scale imply a positive correlation between sales growth and change in efficiency. Alternatively, diminishing marginal returns to new investment imply a negative correlation between sales growth and change in efficiency (because increased sales lead to lower sales prices, which lead to lower turnover).

In line with the discussion at the beginning of this section, growth in real investment should be reflected by the growth component of accruals. In contrast, temporary accounting distortions, such as the overstatement of inventory, should be captured by the efficiency component of accruals. Correlated changes in both growth and distortions will be picked up by the interaction. Note that one potential limitation of these tests is that the growth component of accruals could also be affected by accounting distortions in sales (e.g., overstatement of accounts receivable). Thus, we cannot unambiguously attribute the properties of the growth component to diminishing returns. We can, however, rule out the possibility that properties of the efficiency component are attributable to diminishing returns to new investment. Thus, evidence of a significant role for the efficiency component of accruals in explaining the lower persistence of earnings is consistent with a significant role for temporary accounting distortions.

---

[4] FWY and RSST use average total assets as the deflator. We use lagged net operating assets, as this allows for a parsimonious algebraic decomposition. As a practical matter, however, our empirical results are qualitatively identical if we deflate by either end-of-year *NOA*, average *NOA*, or average total assets.

[5] See the Appendix for the proof.

III. THE EFFECT OF ACCOUNTING DISTORTIONS

We now turn to our numerical examples in order to better understand the impact of accounting distortions on accruals and profitability. Our examples are adapted from related examples in Penman (2001) modeling the effects of permanent accounting distortions. We extend Penman's (2001) examples to consider both permanent and temporary accounting distortions.[6] Recall that the empirical regularity we seek to explain is the positive association between accruals and contemporaneous profitability and the negative association between accruals and subsequent changes in profitability. FWY refer to Penman's (2001) examples to support their claim that this empirical regularity is consistent with the interaction of permanently conservative accounting and growth.[7]

Our numerical examples are based on a generic investment opportunity generating a constant economic return of 10 percent. In order to give the investment opportunity the characteristics of the business operations of a typical firm, we assume that the investment generates sales revenue and requires operating expenditures. The investment is made (in cash) at the end of the current period and generates sales revenues (in cash) at the end of the next period in the amount of 165 percent of the investment. Operating expenses are also incurred (in cash) at the end of the next period in the amount of 55 percent of the investment. These assumptions generate cash flow at the end of the next period equal to 110 percent of the current period investment, to give the 10 percent economic rate of return. The appropriate accounting to capture the underlying economics of this simple investment opportunity should be clear. The benefits to investment made at the end of period t are all realized at the end of period $t+1$. Hence, all investment in period t should be capitalized at the end of period t and expensed at the end of period $t+1$. Such accounting will always result in accounting rates of return that correspond to the economic rate of return of 10 percent (see subsequent examples for details), and we refer to such accounting as "neutral accounting."

In order to study the effect of accounting distortions on accruals and profitability, we introduce "conservative" and "aggressive" accounting alternatives to the neutral accounting described above. Our conservative accounting method only capitalizes 80 percent of the investment made at the end of each period, despite the fact that all the benefits from the investment are received at the end of the subsequent period. An example of such accounting would be a situation where 20 percent of the investment is determined to represent research and development costs, and hence immediately expensed under GAAP. On the other hand, our aggressive accounting method capitalizes 120 percent of the investment made at the end of each period. This is accomplished by capitalizing some of the cash operating costs for that period in addition to the entire investment. An example of such an accounting practice would be the capitalization of maintenance costs that are not expected to generate benefits in subsequent periods. Note that under both the aggressive and conservative accounting methods, we assume that all amounts capitalized in any period are expensed at the end of the subsequent period.[8]

---

[6] We stress that our numerical examples are not intended to provide evidence on the role of diminishing marginal returns to investment in explaining the lower persistence of the accrual component of earnings. Neither Penman (2001) nor FWY use examples to illustrate the impact of diminishing marginal returns to investment. Note that if we construct a model in which we incorporate an assumption of diminishing marginal returns, it is intuitively obvious that returns will diminish, so such a modeling exercise is redundant.

[7] Penman (2001, 561) refers to an example illustrating that growth in real investment interacts with conservative accounting methods to understate profitability in high accrual/investment periods and overstate profitability in subsequent periods. This example is opposite to the empirical regularity that we are trying to explain. Penman (2001, 565) makes the (correct) point that steady-state investment combines with conservative accounting methods to permanently inflate accounting rates of return. This example is also not descriptive of our empirical regularity.

The second dial that we turn in our example is the growth rate in investment. Recall that FWY propose an explanation based on the interaction of real investment growth with permanent accounting distortions. We therefore construct our example to cover seven consecutive periods and we consider how investment growth affects accounting rates of return over these seven periods. We model a positive growth scenario, initially starting with a 0 percent growth rate for two periods, increasing to a 5 percent growth rate for the next period, increasing to 20 percent for one period, then declining to a 5 percent growth rate for one period, followed by a 0 percent growth rate for the final two periods. We select this pattern of growth as it is representative of the sequence of growth rates experienced by a typical "high-growth" firm, with the period of 20 percent growth representing the high growth period (see discussion of Table 3 in Section V). Growth tends to be rapidly mean reverting, so firms with unusually high growth in investment in one period tend to have a declining pattern of growth in surrounding periods (see Nissim and Penman 2001). We also briefly describe the corresponding negative growth scenario that is the mirror image of the positive growth scenario with growth initially starting at 0 percent for two periods, followed by −5 percent for one period, −20 percent for one period, −5 percent for one period, and 0 percent for the final two periods.

Within each of the growth scenarios, we consider five different accounting regimes: the neutral accounting scenario, two aggressive accounting scenarios, and two conservative accounting scenarios. The reason that we have two aggressive and conservative accounting scenarios is that we want to consider the effects of both permanent and temporary accounting distortions. Recall that permanent accounting distortions result from the consistent application of biased accounting methods. These are the types of distortions considered by Penman (2001) and FWY. Temporary accounting distortions, in contrast, result from transitory estimation errors in accruals. These are the types of distortions considered by Dechow and Dichev (2002) and RSST. Our numerical example reports the patterns in profitability (*RNOA*, measured as operating income on lagged operating assets) and the accrual component of profitability (*ACC*, measured as the percentage change in net operating assets) for each of the above scenarios.

Table 1 details our examples. We begin with an initial investment of $100 and then track the financial performance over the subsequent seven periods. We introduce any temporary accounting distortions in the fourth period, labeling this period as event period 0 and other periods relative to this period. The first row of Panel A simply lists the growth rates in investment. The second row lists the dollar amount of investment in each period that results from sequentially applying these growth rates to the initial investment of 100 in period −4. The third and fourth rows show the amounts of operating costs and investment costs that are capitalized in each period. In the neutral accounting scenario, none of the operating costs are capitalized and 100 percent of the investment costs are capitalized in

---

[8] We assume the benefits from the investment and associated amortization of capitalized investment are all realized in the subsequent period. These assumptions are clearly arbitrary and are chosen to be a reasonable but not exhaustive representation. The key takeaways from the examples are robust to assuming that the investment benefits and amortization expenses occur over multiple future periods. The only major qualification is that the accounting distortions reverse over more than one future period. This finding is comforting, since the impact of accrual reversals on earnings also takes more than one period in our empirical data (see Figures 1 and 2).

TABLE 1
Numerical Examples Illustrating the Effect of Accounting Distortions on Profitability and Accruals

Numerical examples illustrating the relation between profitability (*RNOA*) and accruals (*ACC*) for firms experiencing a combination of positive growth and accounting distortions. The numerical examples are based on a simple firm making an investment in period *t* that generates sales in period *t*+1. Sales are 165% of the period *t* investment and operating expenses are 55% of the period *t* investment (both incurred in cash in period *t*+1). In the neutral accounting example, 100% of investment is capitalized in period *t* and expensed in period *t*+1. In the permanently (temporarily) conservative accounting case the firm always (in year 0 only) capitalizes 80% of the investment in year *t* and expenses it in *t*+1. In the permanently (temporarily) aggressive accounting case the firm always (in year 0 only) capitalizes 120% of the investment in year *t* and expenses it in *t*+1. *RNOA* is measured as Operating Income divided by lagged *NOA*. Accruals (*ACC*) are defined as the change in *NOA* divided by lagged *NOA*.

	−4	−3	−2	−1	0	1	2	3
Panel A: Neutral Accounting with Positive Growth								
Growth Rate in Investment		0.0%	0.0%	5.0%	20.0%	5.0%	0.0%	0.0%
Investment	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Capitalized Operating Costs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Capitalized Investment Costs	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Total Capitalized Costs	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Sales	0.00	165.00	165.00	165.00	173.25	207.90	218.30	218.30
Operating Expense	0.00	55.00	55.00	55.00	57.75	69.30	72.77	72.77
Investment Expense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Amortization Expense	0.00	100.00	100.00	100.00	105.00	126.00	132.30	132.30
Operating Income	0.00	10.00	10.00	10.00	10.50	12.60	13.23	13.23
Net Operating Assets (NOA)	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
RNOA		**10.0%**	**10.0%**	**10.0%**	**10.0%**	**10.0%**	**10.0%**	**10.0%**
Accruals (ACC)		**0.0%**	**0.0%**	**5.0%**	**20.0%**	**5.0%**	**0.0%**	**0.0%**
Panel B: Permanently Conservative Accounting with Positive Growth								
Growth Rate in Investment		0.0%	0.0%	5.0%	20.0%	5.0%	0.0%	0.0%
Investment	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Capitalized Operating Costs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Capitalized Investment Costs	80.00	80.00	80.00	84.00	100.80	105.84	105.84	105.84
Total Capitalized Costs	80.00	80.00	80.00	84.00	100.80	105.84	105.84	105.84
Sales	0.00	165.00	165.00	165.00	173.25	207.90	218.30	218.30
Operating Expense	0.00	55.00	55.00	55.00	57.75	69.30	72.77	72.77
Investment Expense	20.00	20.00	20.00	21.00	25.20	26.46	26.46	26.46
Amortization Expense	0.00	80.00	80.00	80.00	84.00	100.80	105.84	105.84
Operating Income	−20.00	10.00	10.00	9.00	6.30	11.34	13.23	13.23
Net Operating Assets (NOA)	80.00	80.00	80.00	84.00	100.80	105.84	105.84	105.84
RNOA		**12.5%**	**12.5%**	**11.3%**	**7.5%**	**11.3%**	**12.5%**	**12.5%**
Accruals (ACC)		**0.0%**	**0.0%**	**5.0%**	**20.0%**	**5.0%**	**0.0%**	**0.0%**

(continued on next page)

TABLE 1 (Continued)

	−4	−3	−2	−1	0	1	2	3
Panel C: Permanently Aggressive Accounting with Positive Growth								
Growth Rate in Investment		0.0%	0.0%	5.0%	20.0%	5.0%	0.0%	0.0%
Investment	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Capitalized Operating Costs	20.00	20.00	20.00	21.00	25.20	26.46	26.46	26.46
Capitalized Investment Costs	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Total Capitalized Costs	120.00	120.00	120.00	126.00	151.20	158.76	158.76	158.76
Sales	0.00	165.00	165.00	165.00	173.25	207.90	218.30	218.30
Operating Expense	−20.00	35.00	35.00	34.00	32.55	42.84	46.31	46.31
Investment Expense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Amortization Expense	0.00	120.00	120.00	120.00	126.00	151.20	158.76	158.76
Operating Income	20.00	10.00	10.00	11.00	14.70	13.86	13.23	13.23
Net Operating Assets (NOA)	120.00	120.00	120.00	126.00	151.20	158.76	158.76	158.76
RNOA		**8.3%**	**8.3%**	**9.2%**	**11.7%**	**9.2%**	**8.3%**	**8.3%**
Accruals (ACC)		**0.0%**	**0.0%**	**5.0%**	**20.0%**	**5.0%**	**0.0%**	**0.0%**
Panel D: Temporarily Conservative Accounting with Positive Growth								
Growth Rate in Investment		0.0%	0.0%	5.0%	20.0%	5.0%	0.0%	0.0%
Investment	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Capitalized Operating Costs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Capitalized Investment Costs	100.00	100.00	100.00	105.00	100.80	132.30	132.30	132.30
Total Capitalized Costs	100.00	100.00	100.00	105.00	100.80	132.30	132.30	132.30
Sales	0.00	165.00	165.00	165.00	173.25	207.90	218.30	218.30
Operating Expense	0.00	55.00	55.00	55.00	57.75	69.30	72.77	72.77
Investment Expense	0.00	0.00	0.00	0.00	25.20	0.00	0.00	0.00
Amortization Expense	0.00	100.00	100.00	100.00	105.00	100.80	132.30	132.30
Operating Income	0.00	10.00	10.00	10.00	−14.70	37.80	13.23	13.23
Net Operating Assets (NOA)	100.00	100.00	100.00	105.00	100.80	132.30	132.30	132.30
RNOA		**10.0%**	**10.0%**	**10.0%**	**−14.0%**	**37.5%**	**10.0%**	**10.0%**
Accruals (ACC)		**0.0%**	**0.0%**	**5.0%**	**−4.0%**	**31.3%**	**0.0%**	**0.0%**
Panel E: Temporarily Aggressive Accounting with Positive Growth								
Growth Rate in Investment		0.0%	0.0%	5.0%	20.0%	5.0%	0.0%	0.0%
Investment	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Capitalized Operating Costs	0.00	0.00	0.00	0.00	25.20	0.00	0.00	0.00
Capitalized Investment Costs	100.00	100.00	100.00	105.00	126.00	132.30	132.30	132.30
Total Capitalized Costs	100.00	100.00	100.00	105.00	151.20	132.30	132.30	132.30
Sales	0.00	165.00	165.00	165.00	173.25	207.90	218.30	218.30
Operating Expense	0.00	55.00	55.00	55.00	32.55	69.30	72.77	72.77
Investment Expense	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Amortization Expense	0.00	100.00	100.00	100.00	105.00	151.20	132.30	132.30
Operating Income	0.00	10.00	10.00	10.00	35.70	−12.60	13.23	13.23
Net Operating Assets (NOA)	100.00	100.00	100.00	105.00	151.20	132.30	132.30	132.30
RNOA		**10.0%**	**10.0%**	**10.0%**	**34.0%**	**-8.3%**	**10.0%**	**10.0%**
Accruals (ACC)		**0.0%**	**0.0%**	**5.0%**	**44.0%**	**−12.5%**	**0.0%**	**0.0%**

each period. The next five rows present the periodic income statements. Sales in each period are equal to 165 percent of the prior period investment and operating expenses are equal to 55 percent of the prior period investment. Investment expense is zero (because 100 percent of investment costs are capitalized) and amortization expense is equal to 100 percent of the prior period investment. Operating income is derived by subtracting expenses from sales. The next row represents the net operating assets at the end of the period and is simply equal to the amount of costs that have been capitalized for that period (remember that all capitalized costs are assumed to be expensed at the end of the subsequent period, so we do not have to consider operating assets from previous periods). The final two rows report the resultant trends in profitability (*RNOA*, measured as operating income on lagged operating assets) and accruals (*ACC*, measured as the percentage change in net operating assets).

Panel A of Table 1 reports the neutral accounting scenario. *RNOA* is constant and equal to the economic rate of return of 10 percent for all periods. This base case illustrates that in the absence of accounting distortions, the accounting rate of return simply equals the economic rate of return. Note also that in this base case scenario, *ACC* is synonymous with the growth rate in real investment.

Panel B of Table 1 introduces permanently conservative accounting into the mix. This example gives us our first look at how growth in investment and accounting distortions interact to produce changes in accounting rates of return. It shows that combining permanently conservative accounting with growth in investment causes the accounting rate of return to deviate from the economic rate of return of 10 percent. With the exception of year 0, the *RNOA* is always higher than the economic rate of return of 10 percent. In year 0, however, the *RNOA* is only 7.5 percent. To understand why this happens, note that conservative accounting can affect both the numerator (i.e., operating income) and the denominator (i.e., net operating assets) of the *RNOA* computation. The denominator effect is to always understate investment, leading to an upward bias in reported *RNOA*. The numerator effect depends on investment growth. With zero growth in investment, there is no net effect on the numerator. This is because the impact of the origination of new conservative accounting exactly offsets the impact of the reversal of old conservative accounting. With positive growth in investment, the numerator is depressed relative to the neutral accounting scenario. This is because the origination of new conservative accounting depresses income by more than the reversal of old conservative accounting boosts income. If the growth rate is greater than the economic rate of return, then the numerator effect will dominate the denominator effect, causing *RNOA* to drop below the economic rate of return (i.e., period 0 of Panel B). Similar effects are illustrated by the examples in Penman (2001, 567 and 571). Finally, note that as in the previous example, *ACC* continues to be synonymous with the growth rate in real investment. This is because permanently conservative accounting understates net operating assets by the same percentage in each period, so the growth rate in net operating assets is not affected.

FWY hypothesize that the lower persistence of the accrual component of earnings is consistent with a combination of growth in investment and conservative accounting methods. However, the patterns in accruals and profitability in Panel B of Table 1 are clearly inconsistent with the empirical regularity described in Section II (i.e., periods with high accruals having high contemporaneous *RNOA*, and lower subsequent *RNOA*). Instead, we see that the increase in accruals in period 0 is associated with low contemporaneous *RNOA* and higher subsequent *RNOA*.

Panel C of Table 1 considers a combination of permanently aggressive accounting with positive growth. Recall from Panel B of Table 1 that permanently conservative accounting combines with positive growth to reduce *RNOA*. For analogous reasons, permanently aggressive accounting combines with positive growth to increase *RNOA*. This happens because the origination of new aggressive accounting increases income by more than the reversal of old aggressive accounting reduces income. This causes both accruals and *RNOA* to be overstated in year 0 and drop back down in year 1. These patterns are consistent with the empirical regularity we are trying to explain. High accruals are associated with high contemporaneous *RNOA* and lower subsequent *RNOA*. Thus, in contrast to the arguments in FWY, permanently aggressive accounting and growth combine to explain the regularity, but permanently conservative accounting and growth do not.

Panel D of Table 1 considers a combination of positive investment growth and temporarily conservative accounting. Neutral accounting is applied for event years −4 through −1, conservative accounting is applied in year 0, and neutral accounting is again applied in years 1 through 3. The application of conservative accounting in year 0 has implications for both year 0 and year 1. Conservative accounting in year 0 deflates year 0 *RNOA* to −14 percent. The conservative accounting also causes accruals to temporarily understate real investment, causing *ACC* to fall to −4 percent. Year 1 sees the reversal of the conservative accounting from period 0, causing income to be overstated. Conservative accounting also causes net operating assets at the beginning of period 1 (i.e., the end of period 0) to be understated. These two effects combine to inflate period 1 *RNOA* to 37.5 percent. *ACC* is also overstated in year 1, because net operating assets are understated at the end of period 0.

The patterns observed in Panel D of Table 1 are broadly consistent with the empirical regularity described in Section II. Temporarily conservative accounting causes low accruals and low contemporaneous *RNOA*, followed by increased accruals and *RNOA* in the subsequent period. Note that the growth rate in investment is not critical to achieving this result. Regardless of the growth rate in investment, temporarily conservative accounting will depress current period *RNOA* and *ACC* and inflate subsequent period *RNOA* and *ACC*. A representative real-world example of such an accounting distortion would be the write-down of inventory or PP&E. Write-downs depress current period profitability and accruals, but inflate future period profitability and accruals.

Finally, Panel E of Table 1 considers the combination of temporarily aggressive accounting with positive growth in investment. This panel represents the flip side of the temporarily conservative accounting in Panel D. Temporarily aggressive accounting causes high accruals and contemporaneous *RNOA*, followed by lower accruals and *RNOA* in the subsequent period. These patterns are also consistent with the empirical regularity described in Section II. Note again that the growth rate in real investment is not required to achieve this result. Even in the absence of growth, temporarily aggressive accounting will inflate current period *RNOA* and *ACC* and deflate subsequent period *RNOA* and *ACC*.

There are three key takeaways from Table 1. First, in contrast to the arguments in FWY, the numerical examples indicate that the interaction of growth in real investment and permanently conservative accounting is inconsistent with the empirical regularity we are trying to explain. In fact, permanently conservative accounting causes patterns in accruals and *RNOA* that are exactly opposite to those observed empirically. Second, consistent with Dechow and Dichev (2002) and RSST, the examples indicate that the empirical results are consistent with temporary accounting distortions. Note that temporary accounting distortions achieve this result even in the absence of growth in real investment. Finally, the examples also indicate that the empirical results are consistent with a combination of growth in real investment and permanently aggressive accounting.[9] Note, however, that our examples ignore the possibility of alternative explanations, such as diminishing marginal returns to investment. The examples also do not help us to discriminate between the competing explanations based on temporary accounting distortions versus growth combined with permanently aggressive accounting distortions. To address these issues, we turn next to our empirical tests.

IV. DATA

Our empirical tests employ data from two sources. Financial statement data are obtained from the Compustat Industrial annual database. Information relating to SEC enforcement actions is found using the Lexis-Nexis Academic Universe product. Our sample period covers all firm years with available data on Compustat for the period 1962–2001. We also exclude all firm-year observations with SIC codes in the range 6000–6999 (financial companies). We eliminate firm-year observations with insufficient data on Compustat to compute the primary financial statement variables used in our tests.[10] These criteria yield a final sample size with non-missing financial statement and returns data of 106,423 firm-year observations. Finally, we identify firm-years that were subject to enforcement actions for alleged GAAP violations associated with earnings overstatements by reading SEC enforcement actions reported on Lexis-Nexis over our sample period. SEC enforcement actions are only available from 1979 onward, so we limit ourselves to this period for the SEC enforcement action tests, reducing this sample to 76,165 firm-year observations. Of this sample, 169 firm-years are subject to SEC enforcement actions for alleged earnings overstatements.

We calculate our measure of accruals (ACC) as $(NOA_t - NOA_{t-1})/NOA_{t-1}$. NOA is the difference between operating assets (OA) and operating liabilities (OL). OA is equal to the sum of current operating assets (COA) and noncurrent operating assets ($NCOA$) and OL is equal to the sum of current operating liabilities (COL) and noncurrent liabilities ($NCOL$). These accrual components are calculated as follows:

COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (STI) (Compustat Item #1).

$NCOA$ = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32).

COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34).

---

[9] In unreported numerical examples we also investigate the interaction of accounting distortions with negative growth in real investment. There are two takeaways. First, temporarily conservative and aggressive accounting continue to explain the empirical regularity in Sloan (1996) and FWY. Second, permanently aggressive accounting interacts with negative growth to explain the empirical regularity. In other words, permanently aggressive accounting interacts with negative growth to temporarily depress accruals and profitability.

[10] Specifically, we require availability of Compustat data items 1, 4, 5, 6, 12, and 181 in both the current and previous year and data item 178 in the current year in order to keep a firm-year observation in the sample. If data items 9, 32 or 34 are missing, then we set them equal to zero rather than eliminating the observation. These data items represent balance sheet items that may not be relevant for some companies (e.g., investments and advances), so we set them to zero rather than needlessly discarding observations. All results are qualitatively similar if we instead estimate each regression using only observations with all data available for that particular regression.

$NCOL$ = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9).

Recall that the decomposition in Equation (4) calls for deflation by lagged NOA.[11] We exclude firm-year observations with negative values of this deflator variable from our analysis. Following previous research, all financial statement ratios are winsorized at $+/-1$ respectively to mitigate the influence of outliers. *RNOA* is calculated as operating income after depreciation (Compustat Item #178) deflated by lagged net operating assets.[12] For our analysis of the persistence in the various components of earnings, we use this same definition of earnings. For our growth versus efficiency decomposition, we compute sales growth (*SG*) as the change in sales (Compustat item #12) scaled by lagged sales and we compute change in asset turnover (ΔAT) as shown in Equation (4).

V. RESULTS

We present our results in three subsections. We begin by presenting descriptive statistics for accruals and related financial variables. These descriptive statistics allow us to examine the correspondence between the empirical properties of accruals and profitability and the properties generated in our examples. We next present tests examining the relation of the growth and efficiency components of accruals with respect to future operating profitability. Finally, we examine the relation between accruals and SEC enforcement actions for alleged overstatements of earnings.

Descriptive Statistics

Panel A of Table 2 provides descriptive statistics for accruals (*ACC*) and related financial variables. The mean and median values for *ACC* are both positive, indicating that the typical firm's operating assets have grown over the sample period. The standard deviations of *SG* and ΔAT are both around 0.3, indicating that they are both important sources of variation in *ACC*. Finally, both the mean and median values of *RNOA* are around 15 percent.[13]

Panel B of Table 2 reports pair-wise correlations for the financial variables. *SG* and ΔAT are strongly positively correlated, indicating that growing firms tend to experience increases in asset turnover, which is consistent with economies of scale. The positive correlation between these two components is indicative of a potentially important role for the interaction component of the sales and efficiency decomposition that we derived in Equation

---

[11] Fairfield et al. (2003b) argue that the lower persistence of accruals relative to cash flows is primarily attributable to growth in the investment base that is not matched by growth in income. In other words, firms with high operating accruals are firms that are growing their operating investments, but earn a lower average return on their increased operating investments, leading to lower persistence in earnings performance. To investigate the importance of this issue for our results, we replicated all of the results in Table 3 using the Fairfield et al. (2003b) approach of measuring our dependent variable with the same deflator as our independent variables. Our results are robust with respect to this procedure. For example, all accrual components continue to have a significantly negative coefficient.

[12] Note that we use a measure of operating income in our analysis of earnings persistence as it is unaffected by nonrecurring components of net income that are reported further down on the income statement. This increases the overall power of our persistence regressions. However, inferences regarding the relative persistence of the different components of earnings are qualitatively similar if we instead use bottom line net income (Compustat item 172).

[13] This is higher than the *RNOA* reported in Nissim and Penman (2001) of 10 percent, but they compute after-tax *RNOA* in their analysis.

TABLE 2
Descriptive Statistics for Total Operating Accruals and Related Financial Variables
(the sample covers 106,423 firm-year observations for the period 1962–2001)

Panel A: Univariate Statistics

	Mean	Std. Dev.	25%	Median	75%
ACC	0.164	0.342	−0.023	0.092	0.280
SG	0.159	0.295	0.009	0.111	0.252
ΔAT	−0.036	0.316	−0.134	0.006	0.121
$RNOA$	0.149	0.338	0.056	0.154	0.279

Panel B: Correlation Matrix—Pearson (above diagonal) and Spearman (below diagonal) (p-values shown below correlations)

	ACC_t	SG_t	ΔAT_t	$RNOA_t$	$RNOA_{t+1}$	$SECAction_t$
ACC_t	—	0.4932 (0.0001)	−0.6206 (0.0001)	0.2034 (0.0001)	0.0315 (0.0001)	0.0250 (0.0001)
SG_t	0.4763 (0.0001)	—	0.2717 (0.0001)	0.1643 (0.0001)	0.0780 (0.0001)	0.0227 (0.0001)
ΔAT_t	−0.5825 (0.0001)	0.2820 (0.0001)	—	−0.0420 (0.0001)	0.0745 (0.0001)	−0.0118 (0.0001)
$RNOA_t$	0.3450 (0.0001)	0.3406 (0.0001)	−0.0878 (0.0001)	—	0.7541 (0.0001)	0.0040 (0.1966)
$RNOA_{t+1}$	0.1015 (0.0001)	0.2039 (0.0001)	0.0685 (0.0001)	0.7357 (0.0001)	—	−0.0175 (0.0001)
$SECAction_t$	0.0224 (0.0001)	0.0188 (0.0001)	−0.0121 (0.0001)	0.0040 (0.1970)	−0.0181 (0.0001)	—

ACC = total operating accruals defined as the change in Net Operating Assets (NOA). It is calculated as $[(NOA_t - NOA_{t-1})/NOA_{t-1}]$;
NOA = Operating Assets (OA) − Operating Liabilities (OL);
OA = the sum of Current Operating Assets (COA) and Noncurrent Operating Assets ($NCOA$);
OL = the sum of Current Operating Liabilities (COL) and Noncurrent Operating Liabilities ($NCOL$);
COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (STI) (Compustat Item #1);
$NCOA$ = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32);
COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34);
$NCOL$ = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9);
SG = sales growth calculated as $[(Sales_t/Sales_{t-1}) - 1]$, where sales is Compustat Item #12;

$$\Delta AT = \left[\left(\frac{Sales_t}{NOA_t}\right) - \left(\frac{Sales_{t-1}}{NOA_{t-1}}\right)\right] \Big/ \frac{Sales_t}{NOA_t};$$

$RNOA$ = return on net operating assets for the current year. It is calculated as income from continuing operations (Compustat Item #178) deflated by beginning net operating assets; and
$SECAction_t$ = an indicator variable equal to 1 if the firm-year observation is subject to a SEC Enforcement Action in year t, and 0 otherwise.

(4). We also see that ACC, SG, and ΔAT are all systematically related to SEC enforcement actions. We will explore these relations more fully in our subsequent regression analyses.

Table 3 and Figure 1 report temporal trends in financial performance for firms with extreme accruals. We first rank firms annually on the magnitude of ACC and allocate them to decile portfolios in ascending order of ACC. We then track the mean (median) ACC, SG,

TABLE 3
Time-Series Pattern in Total Operating Accruals (*ACC*), Sales Growth (*SG*), Changes in Asset Turnover (ΔAT), and Accounting Rates of Return (*RNOA*) for Extreme Accrual Deciles

(the sample consists of 106,423 firm-year observations from 1962–2001)

	Year								
	−4	−3	−2	−1	0	1	2	3	4
Panel A: Low *ACC* Decile									
ACC	0.175	0.173	0.155	0.084	−0.379	0.100	0.130	0.135	0.133
	(0.086)	(0.083)	(0.064)	(0.010)	(−0.330)	(0.024)	(0.063)	(0.068)	(0.067)
SG	0.166	0.159	0.143	0.082	−0.044	0.016	0.087	0.093	0.096
	(0.102)	(0.095)	(0.076)	(0.036)	(−0.051)	(0.011)	(0.061)	(0.067)	(0.069)
ΔAT	−0.050	−0.054	−0.058	−0.054	0.276	−0.102	−0.060	−0.059	−0.059
	(0.001)	(0.000)	(−0.002)	(0.010)	(0.286)	(−0.016)	(0.005)	(0.004)	(−0.003)
RNOA	0.095	0.070	0.028	−0.034	−0.082	−0.019	0.037	0.064	0.081
	(0.111)	(0.097)	(0.068)	(0.027)	(−0.008)	(0.054)	(0.094)	(0.110)	(0.117)
Panel B: High *ACC* Decile									
ACC	0.186	0.197	0.228	0.287	0.863	0.259	0.182	0.152	0.138
	(0.109)	(0.116)	(0.147)	(0.234)	(1.000)	(0.188)	(0.118)	(0.095)	(0.084)
SG	0.204	0.223	0.263	0.332	0.444	0.306	0.168	0.145	0.134
	(0.145)	(0.159)	(0.197)	(0.264)	(0.397)	(0.245)	(0.130)	(0.113)	(0.106)
ΔAT	−0.023	−0.018	−0.012	−0.007	−0.483	−0.009	−0.035	−0.028	−0.023
	(0.024)	(0.030)	(0.037)	(0.029)	(−0.470)	(0.035)	(0.009)	(0.015)	(0.020)
RNOA	0.135	0.146	0.179	0.226	0.235	0.106	0.104	0.109	0.112
	(0.155)	(0.169)	(0.204)	(0.263)	(0.289)	(0.141)	(0.129)	(0.127)	(0.128)

Mean (median) estimates are reported in the first row (second row in parentheses) for each variable.

We sort firm-year observations into ten equal sized portfolios each year. We then track the mean (median) values of *RNOA*, *SG*, ΔAT, and *ACC* for the extreme deciles for four years either side of the ranking year (year 0).

ACC = total operating accruals defined as the change in Net Operating Assets (*NOA*). It is calculated as $[(NOA_t - NOA_{t-1})/NOA_{t-1}]$;

NOA = Operating Assets (*OA*) − Operating Liabilities (*OL*);

OA = the sum of Current Operating Assets (*COA*) and Noncurrent Operating Assets (*NCOA*);

OL = the sum of Current Operating Liabilities (*COL*) and Noncurrent Operating Liabilities (*NCOL*);

COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (*STI*) (Compustat Item #1);

NCOA = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32);

COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34);

NCOL = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9);

SG = sales growth calculated as $[(Sales_t/Sales_{t-1}) - 1]$, where sales is Compustat Item #12;

$$\Delta AT = \left[\left(\frac{Sales_t}{NOA_t}\right) - \left(\frac{Sales_{t-1}}{NOA_{t-1}}\right)\right] \Big/ \frac{Sales_t}{NOA_t};\text{ and}$$

RNOA = return on net operating assets for the current year. It is calculated as income from continuing operations (Compustat Item #178) deflated by beginning net operating assets.

ΔAT, and *RNOA* for the extreme low-*ACC* and high-*ACC* portfolios for the nine-year period centered on the ranking year. Several regularities are immediately obvious. First, accruals are rapidly mean reverting. Second, consistent with FWY and RSST, accounting rates of return track the patterns in accruals. In particular, we see that both the mean and median *RNOA* for the low (high) accrual decile are lowest (highest) in year 0, but then increase (decrease) sharply in year 1. Third, we see that both *SG* and ΔAT closely track the patterns in accruals. Note, however, that the rate of mean reversion is much greater for ΔAT than it is for SG. In fact, as is clear from Figure 1, Panel D, ΔAT shows signs of weak negative serial correlation. This pattern is consistent with the predictions of our temporarily conservative accounting distortion examples from Panels D and E of Table 1, whereby NOA are temporarily misstated in the year of the distortion.

Overall, these patterns in $RNOA$, SG, and ΔAT are consistent with the temporarily conservative and aggressive accounting distortion examples discussed in Section III. The low accrual decile corresponds to temporarily conservative accounting (see Panel D of Table 1) and the high accrual decile corresponds to temporarily aggressive accounting (see Panel E of Table 1). These patterns are also broadly consistent with the interaction of permanently aggressive accounting with growth in real investment. The high accrual decile corresponds to aggressive accounting combined with positive growth (Panel C of Table 1) and the low accrual decile corresponding to aggressive accounting combined with negative growth. In short, these results are consistent with either temporarily aggressive and conservative accounting or the interaction of permanently aggressive accounting with growth. Contradicting the arguments in FWY, these results are inconsistent with the interaction of permanently conservative accounting with growth in real investment (see Panel B of Table 1).

Persistence Results

Table 4 presents our analysis of the persistence of the cash flow and accrual components of earnings. We use the approach adopted in FWY and RSST and previously described in our development of Equation (1). We document the lower persistence of the accrual component of earnings by regressing future $RNOA$ on current $RNOA$ and ACC and testing for a negative coefficient on ACC. Using this specification, the coefficient on ACC represents the difference between the coefficient on the accrual component of $RNOA$ and the coefficient on the cash component of $RNOA$ (see FWY; RSST). Specifically, in Panel A of Table 4 we estimate the following regression:

$$RNOA_{t+1} = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 ACC_t + \upsilon_{t+1}$$

and hypothesize that $\gamma_2 < 0$.

We report results based on the time-series means and t-statistics of annual cross-sectional regressions (Fama and MacBeth 1973) augmented with the Newey and West (1987) correction for autocorrelation. This approach typically generates a conservative estimate of statistical significance (Loughran and Ritter 2000). Results are qualitatively identical (albeit more statistically significant) using pooled OLS regressions. Consistent with the results in FWY and RSST, we find a statistically negative coefficient on ACC. This corroborates the well-known result that the accrual component of earnings is less persistent than the cash flow component of earnings (e.g., Sloan 1996, FWY; RSST).

The regression in Panel B of Table 4 presents the regression results for our growth and efficiency decomposition of ACC:

$$RNOA_{t+1} = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 SG_t - \gamma_3 \Delta AT_t - \gamma_4 (SG_t * \Delta AT_t) + \upsilon_{t+1}. \qquad (5)$$

This decomposition allows us to investigate the diminishing marginal returns to investment explanation developed by FWY. Recall from Section II that diminishing marginal returns to investment will manifest itself in the SG component of accruals. The regressions indicate that the coefficients on $-\Delta AT$ and SG are both significantly negative, with a somewhat

TABLE 4
Time-Series Means and t-Statistics for Coefficients from Annual Cross-Sectional Regressions of Next Year's Accounting Rate of Return on This Year's Accounting Rate of Return and This Year's Total Operating Accruals

(the sample consists of 106,423 firm-year observations from 1962–2001)

Panel A: OLS Regressions for Accruals

$$RNOA_{t+1} = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 ACC_t + \upsilon_{t+1}$$

	γ_0	γ_1	γ_2	Adj. R^2
Coefficient	0.029	0.722		0.5506
(t-statistic)	(5.42)	(71.31)		
	0.041	0.764	−0.131	0.5722
	(8.77)	(67.33)	(−14.85)	

Panel B: OLS Regressions for Growth and Efficiency Decomposition of Accruals

$$RNOA_{t+1} = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 SG_t - \gamma_3 \Delta AT_t - \gamma_4 (SG_t * \Delta AT_t) + \upsilon_{t+1}$$

	γ_0	γ_1	γ_2	γ_3	γ_4	Adj. R^2
Coefficient	0.034	0.730	−0.041			0.5527
(t-statistic)	(7.25)	(69.82)	(−7.24)			
	0.310	0.730		−0.117		0.5667
	(6.35)	(72.29)		(−15.99)		
	0.027	0.728			−0.062	0.5537
	(5.32)	(69.85)			(−4.73)	
	0.042	0.756	−0.102	−0.151	−0.007	0.576
	(10.24)	(61.83)	(−13.65)	(−17.06)	(−0.93)	

Reported coefficient estimates and adjusted R^2s are the means from 40 annual cross-sectional regressions. t-statistics are based on the time-series of the annual regression coefficient estimates using the Fama and MacBeth approach augmented by the Newey and West (1987) correction for autocorrelation.

ACC = total operating accruals defined as the change in Net Operating Assets (NOA). It is calculated as $[(NOA_t - NOA_{t-1})/NOA_{t-1}]$;

NOA = Operating Assets (OA) − Operating Liabilities (OL);

OA = the sum of Current Operating Assets (COA) and Noncurrent Operating Assets ($NCOA$);

OL = the sum of Current Operating Liabilities (COL) and Noncurrent Operating Liabilities ($NCOL$);

COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (STI) (Compustat Item #1);

$NCOA$ = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32);

COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34);

$NCOL$ = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9);

SG = sales growth calculated as $[(Sales_t/Sales_{t-1}) - 1]$, where sales is Compustat Item #12;

$$\Delta AT = \left[\left(\frac{Sales_t}{NOA_t}\right) - \left(\frac{Sales_{t-1}}{NOA_{t-1}}\right)\right] \Big/ \frac{Sales_t}{NOA_t}; \text{ and}$$

$RNOA$ = return on net operating assets for the current year. It is calculated as income from continuing operations (Compustat Item #178) deflated by beginning net operating assets.

larger and more statistically significant coefficient on $-\Delta AT$.[14] The negative coefficient on $-\Delta AT$ is consistent with accounting distortions providing a significant contribution to the lower persistence of accruals.[15] Note also that since sales are generally based on accrual accounting estimates (e.g., accounts receivable), the negative coefficient on sales growth may be indicative of either diminishing returns to new investment or accounting distortions in sales. In summary, the evidence from the growth versus efficiency regressions is suggestive of a significant role for accounting distortions in explaining the lower persistence of the accrual component of earnings, though it does not rule out a supplementary role for diminishing returns to new investment.

---

[14] Unreported tests indicate that the coefficient on $-\Delta AT$ is significantly more negative than the coefficient on SG at the 0.01 level.

SEC Enforcement Action Results

Tables 5 and 6 present our final set of tests, investigating the relation between accruals and SEC enforcement actions. The purpose of these tests is to examine the role of managerial discretion in explaining the lower persistence of the accrual component of earnings. Finding a positive contemporaneous relation between accruals and SEC enforcement actions would be consistent with the presence of aggressive distortions in accruals that are attributable to managerial discretion. Note that these tests provide evidence of a special case of the transitory accounting estimation error explanation provided in Dechow and Dichev (2002) and RSST. Specifically, following conjectures by Xie (2001) and Chan et al. (2005), these tests investigate whether the transitory distortions in accruals are attributable to intentional managerial manipulation.

Our dependent variable, *SECAction*, takes the value of 1 for firm-years that the SEC subsequently alleges earnings were overstated in violation of GAAP, and 0 otherwise. Due to the dichotomous nature of the dependent variable, we use logistic regression analysis. The regression in Panel A of Table 5 confirms the existence of a significant positive relation between total operating accruals and the likelihood of an SEC enforcement for overstated earnings (see also, Bradshaw et al. 2001; Dechow et al. 1996).

Panel B of Table 5 employs the RSST decomposition of total operating accruals into working capital accruals and noncurrent operating accruals. Recall that both FWY and RSST show that each of these two accrual components lead to lower earnings persistence. FWY, however, argue that the lower persistence of the noncurrent operating component of accruals is unlikely to be attributable to managerial discretion. In contrast to the arguments in FWY, we find that both components of accruals have a strong positive relation with SEC enforcement actions, suggesting that noncurrent operating accruals also reflect managerial manipulation. These results suggest that both components of accruals are characterized by temporarily aggressive accounting distortions that are attributable to managerial discretion.

Panel C of Table 5 report results from regressions of SEC enforcement actions on the growth and efficiency components of accruals. Both components are significantly related to SEC enforcement actions. The significance of sales growth is puzzling. A potential explanation is that sales growth reflects temporary accounting distortions arising from revenue manipulation. Consistent with this explanation, prior research has documented that revenue manipulation is a leading technique of earnings manipulation (see Dechow et al. 1995; Dechow and Schrand 2004). Assuming that revenue manipulation explains the significance of sales growth, the negative coefficient on *SG* in Table 4 may also be driven by accounting distortions as opposed to diminishing marginal returns to investment. There are,

---

[15] The sales growth and efficiency measures potentially contain measurement error for firms engaging in merger or acquisition activity during the sample year. As a robustness check, we eliminate all observations with an increase in goodwill from prior year and re-estimated the regressions (reducing the sample size to 96,116 observations). The results are qualitatively similar.

TABLE 5
Logistic Regressions of SEC Enforcement Actions on Total Operating Accruals

(the sample consists of 76,165 firm-year observations from 1979 to 2001)

Panel A: Logistic Regressions for Total Operating Accruals

$$SECAction_t = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 ACC_t + v_t$$

Regression Coefficient	γ_0	γ_1	γ_2	Model Chi-Square	p-value
Estimated value	−6.49	0.21	1.28	51.77	(0.0001)
(p-value)	(0.0001)	(0.2574)	(0.0001)		

Panel B: Logistic Regressions for Working Capital and Noncurrent Operating Components of Accruals

$$SECAction_t = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 \Delta WC_t + \gamma_3 \Delta NCO_t + v_t$$

Regression Coefficient	γ_0	γ_1	γ_2	γ_3	Model Chi-Square	p-value
Estimated value	−6.29	0.26	1.24		32.59	(0.0001)
(p-value)	(0.0001)	(0.1964)	(0.0001)			
	−6.36	0.35		1.05	30.73	(0.0001)
	(0.0001)	(0.0708)		(0.0001)		
	−6.39	0.21	0.94	0.84	50.43	(0.0001)
	(0.0001)	(0.2636)	(0.0001)	(0.0002)		

Panel C: Logistic Regressions for Growth and Efficiency Components of Accruals

$$SECAction_t = \gamma_0 + \gamma_1 RNOA_t + \gamma_2 SG_t - \gamma_3 \Delta AT_t - \gamma_4 (SG_t * \Delta AT_t) + v_t$$

Regression Coefficient	γ_0	γ_1	γ_2	γ_3	γ_4	Model Chi-Square	p-value
Estimated value	−6.49	0.33	1.36			47.06	(0.0001)
(p-value)	(0.0001)	(0.0828)	(0.0001)				
	−6.23	0.44		0.65		14.81	(0.0006)
	(0.0001)	(0.0339)		(0.0019)			
	−6.18	0.34			1.29	24.70	(0.0001)
	(0.0001)	(0.1039)			(0.0001)		
	−6.52	0.22	1.40	0.60	0.27	73.22	(0.0001)
	(0.0001)	(0.2237)	(0.0001)	(0.0333)	(0.4129)		

We identified firms subject to a SEC enforcement action via a reading of the SEC Accounting and Auditing Enforcement Releases from 1979–2001.

p-values are presented below the coefficients.

$SECAction_t$ = an indicator variable equal to 1 if the firm-year observation is subject to a SEC Enforcement Action in year t, and 0 otherwise;

$RNOA$ = return on net operating assets for the current year. It is calculated as income from continuing operations (Compustat Item #178) deflated by beginning net operating assets;

ACC = total operating accruals defined as the change in Net Operating Assets (NOA). It is calculated as $[(NOA_t - NOA_{t-1})/NOA_{t-1}]$;

NOA = Operating Assets (OA) − Operating Liabilities (OL);

OA = the sum of Current Operating Assets (COA) and Noncurrent Operating Assets ($NCOA$);

OL = the sum of Current Operating Liabilities (COL) and Noncurrent Operating Liabilities ($NCOL$);

(*continued on next page*)

TABLE 5 (Continued)

COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (*STI*) (Compustat Item #1);
NCOA = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32);
COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34);
NCOL = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9);
Change in working capital (ΔWC) = $[(WC_t - WC_{t-1})/NOA_{t-1}$, where $WC = (COA - COL)$;
Change in noncurrent operating accruals (ΔNCO) = $[(NCO_t - NCO_{t-1})/NOA_{t-1}$, where $NCO = (NCOA - NCOL)$;
SG = sales growth calculated as $[(Sales_t/Sales_{t-1}) - 1]$, where sales is Compustat Item #12; and

$$\Delta AT = \left[\left(\frac{Sales_t}{NOA_t}\right) - \left(\frac{Sales_{t-1}}{NOA_{t-1}}\right)\right] \bigg/ \frac{Sales_t}{NOA_t}.$$

TABLE 6

Regressions of Total Operating Accrual Ranks on Current and Lagged SEC Enforcement Actions

(the sample consists of 76,165 firm-year observations from 1979 to 2001)

$$Rank\ of\ ACC_t = \gamma_0 + \gamma_1 SECAction_t + \gamma_2 SECAction_{t-1} + \gamma_3 SECAction_{t-2} + \gamma_4 SECAction_{t-3} + \upsilon_t$$

		γ_0	γ_1	γ_2	γ_3	γ_4	Adj. R^2
Coefficient (t-statistic)	Model 1:	0.478	0.155				0.0005
		(394.67)	(5.89)				
	Model 2:	0.479	0.192	−0.078			0.0006
		(394.62)	(6.36)	(−2.51)			
	Model 3:	0.467	0.169	0.023	−0.165		0.0009
		(364.59)	(5.18)	(0.62)	(−4.71)		
	Model 4:	0.460	0.165	0.027	−0.119	−0.081	0.0009
		(339.47)	(4.59)	(0.69)	(−2.90)	(−2.13)	

We assign firms into ten-equal-sized portfolios each year based on the magnitude of *ACC*. The rank is standardized to range from 0 to 1. We use this rank as our dependent variable in the regression analysis above. Results are similar if we instead use the raw accrual measure.

$SECAction_t$ = an indicator variable equal to 1 if the firm-year observation is subject to a SEC Enforcement Action in year *t*, and 0 otherwise. We identified firms subject to a SEC enforcement actions via a reading of the SEC Accounting and Auditing Enforcement Releases from 1979–2001;
ACC = total operating accruals defined as the change in Net Operating Assets (*NOA*). It is calculated as $[(NOA_t - NOA_{t-1})/NOA_{t-1}]$;
NOA = Operating Assets (*OA*) − Operating Liabilities (*OL*);
OA = the sum of Current Operating Assets (*COA*) and Noncurrent Operating Assets (*NCOA*);
OL = the sum of Current Operating Liabilities (*COL*) and Noncurrent Operating Liabilities (*NCOL*);
COA = Current Assets (Compustat Item #4) − Cash and Short-Term Investments (*STI*) (Compustat Item #1);
NCOA = Total Assets (Compustat Item #6) − Current Assets (Compustat Item #4) − Investments and Advances (Compustat Item #32);
COL = Current Liabilities (Compustat Item #5) − Debt in Current Liabilities (Compustat Item #34); and
NCOL = Total Liabilities (Compustat Item #181) − Current Liabilities (Compustat Item #5) − Long-Term Debt (Compustat Item #9).

---

however, two alternative explanations for this result. First, the SEC may disproportionately target high-growth firms. Such firms are more likely to be raising new capital, thus making them more likely targets for SEC enforcement actions. Second, recall from the numerical example in Section III that permanently aggressive accounting combined with positive growth can cause earnings to be temporarily overstated. Thus, the SEC may only target cases of permanently aggressive accounting in the presence of positive sales growth, because this is the only situation where earnings are overstated.

We can use the SEC enforcement action data to investigate one additional implication of the accounting distortion explanation. Recall from the numerical examples in Table 1 that the negative relation between accruals and future profitability is consistent with both temporary accounting distortions and the interaction of permanently aggressive accounting with real investment growth. One difference between these two cases is that temporarily aggressive accounting distortions involve the understatement of accruals in subsequent periods (see Panel E of Table 1). The interaction of permanently aggressive accounting with growth, in contrast, does not involve the understatement of accruals in future periods (see Panel C of Table 1). To discriminate between these two explanations, we therefore investigate the relation between accruals and *lagged* SEC enforcement actions (in other words, we investigate the relation between SEC enforcement actions and ***subsequent*** accruals). If temporarily aggressive accounting distortions drive the high accruals at the time of SEC enforcement actions, then the subsequent reversal of these distortions should lead to a negative relation between accruals and *lagged* SEC enforcement actions. Note that all accrual reversals take place in the very next period in our examples. In practice, however, accrual reversals may take longer than one year. We therefore test for this possibility in Table 6 by examining up to three years of lagged SEC enforcement actions.

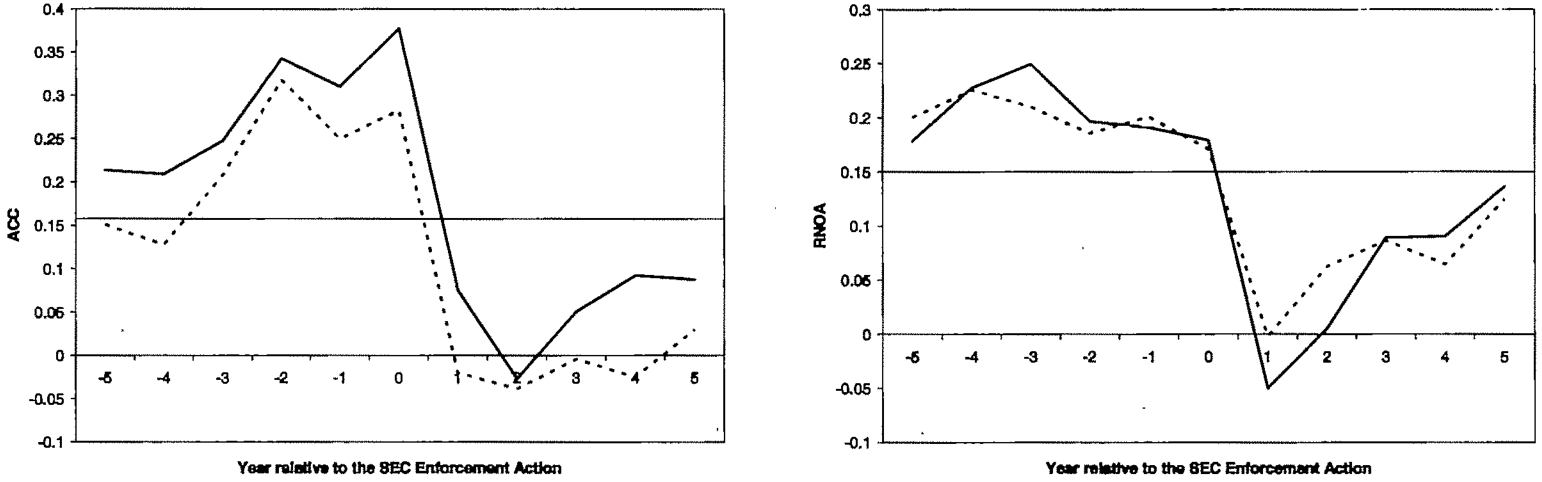
Consistent with the results in Table 5, Table 6 documents a strong positive contemporaneous relation between accruals and SEC enforcement actions (the coefficient on γ_1 is always statistically positive). More importantly, the addition of lagged accruals to the regression reveals evidence of significantly negative coefficients, particularly at the second and third lags. This evidence is consistent with the existence of temporarily aggressive accounting distortions at the time of SEC enforcement actions that take two to three years to reverse. It suggests that, at least for the SEC enforcement action subsample, the properties of accruals are attributable to temporarily aggressive accounting rather than the interaction of permanently aggressive accounting with growth in real investment.

The results from our SEC enforcement action tests are summarized in Figure 2. This figure plots *ACC*, *RNOA*, *SG*, and ΔAT for the SEC sample over the 11-year window centered on the alleged earnings overstatement. Panel A plots the mean and median values of *ACC* for the SEC sample. The solid line across the middle of the graph plots the average *ACC* for our entire sample of 0.164 as a visual benchmark. The graph clearly shows that accruals are unusually high in the year of the alleged earnings overstatement and the two preceding years. Accruals then post a sharp decline in the year immediately following the overstatement and remain unusually low for the next several years. Panel B plots the mean and median values of *RNOA* for the SEC sample. This graph shows that the sharp reduction in accruals between year 0 and year 1 causes a corresponding sharp reduction in *RNOA*. *RNOA* is unusually high in the years leading up to the SEC enforcement action and unusually low thereafter. Panel C demonstrates that the *SG* component of accruals is temporarily high in the year of the alleged earnings overstatement and temporarily low in the following years. This result is consistent with revenue-based earnings manipulations that temporarily inflate sales growth. Panel D demonstrates that the ΔAT component of accruals is temporarily low in the year of the alleged earnings overstatement and temporarily high in the following years. This result is consistent with temporarily aggressive accounting distortions at the time of SEC enforcement actions that reverse in subsequent years.

FIGURE 2

Event-Time Plots of Accruals (*ACC*), Accounting Rate of Return (*RNOA*), Sales Growth (*SG*), and Change in Asset Turnover (*ΔAT*) for Firms Subject to SEC Enforcement Actions

Year 0 represents the year that a firm was subject to an SEC enforcement action for an alleged earnings overstatement. The solid (dashed) line reflects the mean (median) of the respective variable.

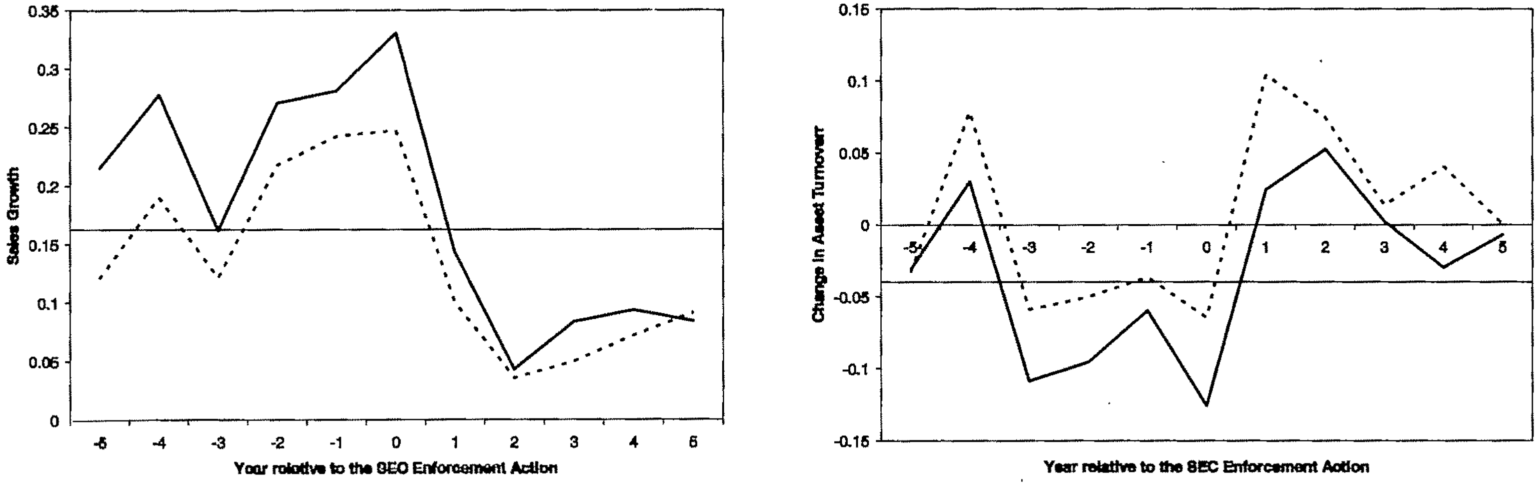


Panel A: Accruals (*ACC*)

Panel B: Accounting Rates of Return (*RNOA*)

(*continued on next page*)

FIGURE 2 (continued)



Panel C: Sales Growth (*SG*)

Panel D: Change in Asset Turnover (Δ*AT*)

VI. CONCLUSION

This paper examines the role of accounting distortions in explaining the lower persistence of the accrual component of earnings. A number of recent papers have argued that the lower persistence of the accrual component of earnings is more likely to be attributable to diminishing marginal returns to new investment and other "growth-based" explanations rather than to accounting-based explanations. In contrast, our analysis and empirical evidence suggests that temporary accounting distortions play an important role. At the same time, we are unable to rule out a supplementary role for growth-based explanations, such as diminishing marginal returns to new investment. We conclude that temporary accounting distortions are a significant contributing factor to the lower persistence of the accrual component of earnings. Our analysis of SEC enforcement actions also indicates that some of these accounting distortions are attributable to intentional managerial manipulation of accruals.

Our findings reinforce the importance of accounting analysis in the evaluation of financial performance. Accrual accounting seeks to provide more relevant information to investors. The increase in relevance, however, comes at the cost of reduced reliability. Reduced reliability manifests itself in the form of transitory estimation error in accruals and increased opportunities for the opportunistic use of managerial accounting discretion in setting accruals. Our evidence indicates that the lower persistence of the accrual component of earnings is a direct manifestation of this crucial trade-off between relevance and reliability. Our findings suggest that the current push by accounting standard setters toward "fair value" accounting is likely to increase the frequency and magnitude of accounting distortions, leading to a reduction in the persistence and reliability of earnings.

Future research could further examine the impact of this trade-off on the properties of accounting numbers. A detailed examination of specific accruals categories (e.g., inventory, receivables) could provide detailed corroborating evidence on the role of temporary accounting distortions. Additional topics for future research include: (1) the determination of which accruals are most subject to accounting distortions; (2) the detailed documentation of the origination and reversal of such distortions; and (2) the identification of external factors facilitating accounting distortions, such as subjective accounting standards and weak corporate governance.

APPENDIX
Proof of Growth and Efficiency Decomposition in Equation (4)

$$\text{Total Net Accruals} = \Delta NOA_t / NOA_{t-1}$$

$$= \underbrace{\Delta \text{Sales}_t / Sales_{t-1}}_{\substack{\text{Sales}\\ \text{Growth}}} - \underbrace{\Delta AT_t / AT_t}_{\substack{\text{Change in}\\ \text{Efficiency}}} - \underbrace{(\Delta Sales_t / Sales_{t-1})*(\Delta AT_t / AT_t)}_{\text{Interaction}} \qquad (4)$$

The RHS of this expression can be reduced to the LHS as follows:

$$\frac{S_t - S_{t-1}}{S_{t-1}} - \frac{\frac{S_t}{NOA_t} - \frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} - \frac{S_t - S_{t-1}}{S_{t-1}} * \frac{\frac{S_t}{NOA_t} - \frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}}$$

$$= \frac{S_t - S_{t-1}}{S_{t-1}} - \left(\frac{\frac{S_t}{NOA_t}}{\frac{S_t}{NOA_t}} - \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} \right) - \frac{S_t - S_{t-1}}{S_{t-1}} * \left(\frac{\frac{S_t}{NOA_t}}{\frac{S_t}{NOA_t}} - \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} \right)$$

$$= \frac{S_t - S_{t-1}}{S_{t-1}} - 1 + \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} - \frac{S_t - S_{t-1}}{S_{t-1}} * \left(1 - \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} \right)$$

$$= \frac{S_t - S_{t-1}}{S_{t-1}} - 1 + \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}} - \frac{S_t - S_{t-1}}{S_{t-1}} + \frac{S_t - S_{t-1}}{S_{t-1}} * \frac{\frac{S_{t-1}}{NOA_{t-1}}}{\frac{S_t}{NOA_t}}$$

$$= -1 + \frac{S_{t-1}}{NOA_{t-1}} * \frac{NOA_t}{S_t} + \frac{S_t - S_{t-1}}{S_{t-1}} * \frac{S_{t-1}}{NOA_{t-1}} * \frac{NOA_t}{S_t}$$

$$= -1 + \frac{S_{t-1}}{NOA_{t-1}} * \frac{NOA_t}{S_t} * \frac{S_{t-1}}{S_{t-1}} + \frac{S_t - S_{t-1}}{S_{t-1}} * \frac{S_{t-1}}{NOA_{t-1}} * \frac{NOA_t}{S_t}$$

$$= -1 + \frac{S_{t-1}NOA_tS_{t-1}}{NOA_{t-1}S_tS_{t-1}} + \frac{S_tS_{t-1}NOA_t}{NOA_{t-1}S_tS_{t-1}} - \frac{S_{t-1}S_{t-1}NOA_t}{NOA_{t-1}S_tS_{t-1}}$$

$$= -1 + \frac{S_tS_{t-1}NOA_t}{NOA_{t-1}S_tS_{t-1}}$$

$$= -1 + \frac{NOA_t}{NOA_{t-1}}$$

$$= \frac{NOA_t - NOA_{t-1}}{NOA_{t-1}}$$

REFERENCES

Anderson, C., and L. Garcia-Feijoo. 2006. Empirical evidence on capital investment, growth options, and security returns. *Journal of Finance* 61: 171–194.

Bradshaw, M., S. Richardson, and R. Sloan. (2001). Do analysts and auditors understand information in accruals? *Journal of Accounting Research* 39: 45–74.

Chan, K., L. Chan, N. Jegadeesh, and J. Lakonishok. 2006. Earnings quality and stock returns. *Journal of Business* 79: 1041–1082.

Cooper, M., H. Gulen, and M. Schill. 2005. What best explains the cross-section of stock returns? Exploring the asset growth effect. Working paper, Purdue University.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18: 3–42.

———, R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–226.

———, ———, and ———. 1996. Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research* 13: 1–36.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 35–59.

———, and C. Schrand. 2004. *Earnings Quality*. Charlottesville, VA: CFA Institute.

Fairfield, P., J. Whisenant, and T. Yohn. 2003a. Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review* 78: 353–371.

———, ———, and ———. 2003b. The differential persistence of accruals and cash flows for future operating income versus future profitability. *The Review of Accounting Studies* 8: 221–243.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return and equilibrium: Empirical tests. *The Journal of Political Economy* 81: 607–636.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7: 85–107.

Jansen, I., and T. Yohn. 2003. Are changes in asset turnover informative about earnings management? Working paper, Georgetown University.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.

Khan, M. 2005. Are accruals really mispriced? Evidence from tests of an intertemporal capital asset pricing model. Working paper, Massachusetts Institute of Technology.

Loughran, T., and J. R. Ritter. 2000. Uniformly least powerful tests of market efficiency. *Journal of Financial Economics* 55: 361–389.

Nelson, M., J. Elliot, and R. Tarpley. 2003. How are earnings managed? Examples from auditors. *Accounting Horizons* (Supplement): 17–35.

Newey, W., and K. West. 1987. A simple, positive, semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–8.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* 6: 109–154.

Penman, S. 2001. *Financial Statement Analysis and Equity Valuation*. New York, NY: McGraw Hill.

Richardson, S., R. Sloan, M. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39: 437–485.

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71: 289–315.

Stigler, G. 1963. *Capital and Rates of Return in Manufacturing Industries*. Princeton, NJ: Princeton University Press.

Thomas, J., and H. Zhang. 2001. Inventory changes and future returns. *The Review of Accounting Studies* 7: 163–187.

Titman, S., K. C. Wei, and F. Xie. 2004. Capital investments and stock returns. *Journal of Financial and Quantitative Analysis* 39: 677–700.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76: 357–373.

Zach, Z. 2005. Inside the accrual anomaly. Working paper, Washington University.

Editorial Data

The following table contains information about turnaround time for manuscripts (including revisions) on which editorial decisions were made in the 12-month period ended December 31, 2005. Turnaround time is the number of days between the date that the manuscript was received *and the date of the editor's letter to the author(s).*

	Number of Manuscripts	Percent	Cumulative Number	Cumulative Percent
0 ≤ Days ≤ 30	84	18.06	84	18.06
31 ≤ Days ≤ 60	183	39.36	267	57.42
61 ≤ Days ≤ 90	140	30.11	407	87.53
91 ≤ Days ≤ 120	49	10.54	456	98.07
121 ≤ Days	09	1.93	465	100.00
Total	465	100.00		

The mean review time was 55.93; the median review time was 77 days.

New Submissions by calendar year (January 1 to December 31)

1992	241
1993	234
1994	231
1995	195
1996	230
1997	195
1998	196
1999	239
2000	260
2001	260*
2002	324
2003	327
2004	307
2005	389

* 2001—in addition to the 260 regular submissions, 68 MSs were submitted to the *TAR* Quality of Earnings Conference.

The acceptance rate (defined as number of MSs accepted divided by the number of new submissions) over the past 5 years is between 12–16% per year.

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and associate editors are responsible for providing critically constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation. Associate editors also make editorial recommendations to the editor.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be printed in 12-point font on one side of 8½ × 11" good quality paper and be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins of at least one inch from top, bottom, and sides should facilitate editing and duplication.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, the author's name, title and affiliation, email address, any acknowledgments, and a footnote indicating whether the author would be willing to share the data (see last paragraph in this statement).

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years*. All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a well-presented analysis; re-form*. See *Webster's* for correct usage.

Keywords: The abstract is to be followed by four keywords that will assist in indexing the paper.

ABSTRACT/INTRODUCTION

An Abstract of about 100 words should be presented on a separate page immediately preceding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate by marginal notation where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Work cited should use the "author-date system" keyed to a list of works in the reference list (see below). Authors should make an effort to include the relevant page numbers in the cited works.

1. In the text, works are cited as follows: authors' last name and date, without comma, in parentheses: for example, (Jones 1987); with two authors: (Jones and Freeman 1973); with more than two: (Jones et al. 1985); with more than one source cited together (Jones 1987; Freeman 1986); with two or more works by one author: (Jones 1985, 1987).
2. Unless confusion would result, do not use "p." or "pp." before page numbers: for example, (Jones 1987, 115).
3. When the reference list contains more than one work of an author published in the same year, the suffix a, b, etc. follows the date in the text citation: for example, (Jones 1987a) or (Jones 1987a; Freeman 1985b).
4. If an author's name is mentioned in the text, it need not be repeated in the citation; for example, "Jones (1987, 115) says ..."
5. Citations to institutional works should use acronyms or short titles where practicable; for example, (AAA ASOBAT 1966); (AICPA Cohen Commission Report 1977). Where brief, the full title of an institutional work might be shown in a citation: for example, (ICAEW *The Corporate Report* 1975).
6. If the manuscript refers to statutes, legal treatises or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a list of references containing only those works cited. Each entry should contain all data necessary for unambiguous identification. With the author-date system, use the following format recommended by *The Chicago Manual*:

1. Arrange citations in alphabetical order according to surname of the first author or the name of the institution responsible for the citation.
2. Use author's initials instead of proper names.
3. Date of publication should be placed immediately after author's name.
4. Titles of journals should not be abbreviated.
5. Multiple works by the same author(s) in the same year are distinguished by letters after the date.
6. Inclusive page numbers are treated as recommended in *Chicago Manual* section 17.168.B.

Sample entries are as follows:

American Accounting Association, Committee on Concepts and Standards for External Financial Reports. 1977. *Statement on Accounting Theory and Theory Acceptance*. Sarasota, FL: AAA.

Demski, J. S., and D. E. M. Sappington. 1989. Hierarchical structure and responsibility accounting. *Journal of Accounting Research* 27 (Spring): 40–58.

Dye, R., B. Balachandran, and R. Magee. 1989. Contingent fees for audit firms. Working paper, Northwestern University, Evanston, IL.

Fabozzi, F., and I. Pollack, eds. 1987. *The Handbook of Fixed Income Securities*. 2nd edition. Homewood, IL: Dow Jones-Irwin.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge, U.K.: Cambridge University Press.

Porcano, T. M. 1984a. Distributive justice and tax policy. *The Accounting Review* 59 (October): 619–636.

———. 1984b. The perceived effects of tax policy on corporate investment intentions. *The Journal of the American Taxation Association* 6 (Fall): 7–19.

Shaw, W. H. 1985. Empirical evidence on the market impact of the safe harbor leasing law. Ph.D. dissertation, The University of Texas at Austin.

Sherman, T. M., ed. 1984. *Conceptual Framework for Financial Accounting*. Cambridge, MA: Harvard Business School.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be consecutively numbered throughout the manuscript with superscript Arabic numerals. Footnote text should be doubled-spaced and placed at the end of the article.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state that the work is not submitted or published elsewhere.
2. In the case of manuscripts reporting on field surveys or experiments, four copies of the instrument (questionnaire, case, interview plan, or the like) should be submitted. Information that might identify the author(s) must be deleted from the instrument.
3. MS should be submitted via email as Microsoft® Word or PDF files to the *TAR* office at TARUAZ@email.arizona.edu. A submission letter, a separate cover page, and the MS containing no author identification should be emailed as separate files. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Professor Dan Dhaliwal, Senior Editor, *The Accounting Review*, The University of Arizona, Department of Accounting, McClelland Hall 301, 1130 E. Helen Street, Tucson, AZ 85721-0108. The submission fee in U.S. funds is $125.00 for members and $200.00 for nonmembers of the AAA. Effective January 1990, the submission fee is nonrefundable.
4. The author should retain a copy of the paper.
5. Revisions must be submitted within 12 months from request, otherwise they will be considered new submissions.

COMMENTS

Comments on articles previously published in *The Accounting Review* will be reviewed (anonymously) by two reviewers in sequence. The first reviewer will be the author of the original article being subjected to critique. If substance permits, a suitably revised comment will be sent to a second reviewer to determine its publishability in *The Accounting Review*. If a comment is accepted for publication, the original author will be invited to reply. All other editorial requirements, as enumerated above, also apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the *Review* will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy has been adopted by the Executive Committee in its April 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

MAY 2006 PLACEMENT ADS

*The deadline for information to be included in the "Placement Information" section is two months prior to the date of publication. The **Review** is published in January, April, July, and October. Placement advertising and replies should be sent to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, ATTENTION: ADVERTISING DEPT. All replies to box numbers should be on separate sheets to facilitate forwarding.*

301 ROBERT MORRIS UNIVERSITY, Department of Accounting and Taxation, invites applications for anticipated faculty positions beginning August 2006. Candidates with teaching and research interests in all areas of accounting will be considered. Candidates should either possess a Ph.D. or be nearing completion of their degree program. Appointment will be at the Assistant, Associate, or Full Professor level, depending on the applicant's experience and accomplishments. Located near Pittsburgh, Robert Morris University is a private institution, enrolling approximately 5,000 students. Teaching is mostly at the undergraduate level, although opportunities to teach graduate courses are available. Creating a high-quality, innovative learning environment is the Department's top priority. Scholarly activity is encouraged and supported. Salary and benefits are competitive. Candidates should submit a letter of application, *curriculum vitae*, and names and addresses for three references to: Professor James E. Rebele, School of Business, Robert Morris University, 6001 University Boulevard, Moon Township, PA 15108. Applications may be submitted electronically to: rebele@rmu.edu. Robert Morris University is an Equal Opportunity Employer.

302 UNIVERSITY OF NOTRE DAME, Department of Accountancy, invites applications for position openings at the Assistant and Associate Professor levels. All research areas will be considered. Applicants for Assistant Professor positions should expect to complete doctoral degree requirements before the start of the contract period. All candidates should have a strong commitment to scholarly research and excellence in teaching. Salary is very competitive and research support is excellent. Please send *curriculum vitae* and an example of scholarly work to: Professor Tom Schaefer, Chair, Department of Accountancy, Room 102D, Mendoza College of Business Administration, University of Notre Dame, Notre Dame, IN 46556-5646. The University of Notre Dame is an Equal Opportunity/Affirmative Action Employer.

303 OREGON STATE UNIVERSITY invites applicants for a tenure-track position in Accounting at the rank of Assistant, Associate, or Full to begin September 2006, subject to funding. Teaching duties will be in auditing, financial, and/or tax. Responsibilities include research, quality teaching, and professional activities in the field of accounting. Candidates should hold the appropriate doctorate. Professional certification is desirable. Salary is competitive, fringe benefits are excellent. Oregon State University is located in Corvallis, which is rated the second-best small city in America. Interested applicants should submit a letter of interest, current *vita*, recent teaching evaluations, evidence of scholarly work, and a list of three references to: Dr. Roger Graham, College of Business, 200 Bexell Hall, Oregon State University, Corvallis, Oregon 97331-2603. For full consideration apply by February 28, 2006. OSU is an Equal Opportunity Employer and seeks diversity.

304 NORTHEASTERN ILLINOIS UNIVERSITY, College of Business and Management, invites applications for Ph.D./D.B.A. with specialization in Accounting from an AACSB-accredited College of Business program in one or both of the following areas: Accounting Principles and Financial Accounting. The position is flexible to accommodate other teaching and research interests. One tenure-track position at the Assistant/Associate Professor rank beginning 8/06, is anticipated for teaching undergraduate and graduate courses. Potential for excellent teaching and research. Selection process begins on January 15, 2006 and continues until position is filled. Send application letter, resume with email address, transcripts, and names of three references to: Dr. C. T. Chen, Chair, Department of Accounting, Business Law and Finance, Northeastern Illinois University, 5500 N. St. Louis Avenue, Chicago, IL 60625. Fax: 773-442-6110. Email: C-Chen3@neiu.edu. AA/EOE.

305 UNIVERSITY OF TENNESSEE AT CHATTANOOGA, College of Business Administration, invites applications for an expected tenure-track Assistant or Associate Professor position, primarily teaching managerial accounting in the M.B.A. program. Candidates should have a Ph.D. or D.B.A. in Accounting or have an estimated completion date of August 2006. Candidates must be committed to excellence in teaching. We are especially interested in faculty who want to combine teaching and research with a close working relationship with the Chattanooga community. Salary and benefits are AACSB competitive. Review of applicants will begin on February 1 and continue until the position is filled. For information, contact: Dr. Rick Turpin, University of Tennessee at Chattanooga, College of Business Administration, Department of Accounting and Finance, Dept. 6206, 615 McCallie Avenue, Chattanooga, TN 37403. Telephone: (423) 425-4664; Email: Rick-Turpin@utc.edu. Minority applicants are strongly encouraged to apply. The University of Tennessee at Chattanooga is an Equal Employment Opportunity, Affirmative Action, Title VI and IX, Section 504, ADA Institution.

306 McGILL UNIVERSITY, Desautels Faculty of Management, has tenure-track positions available at the Assistant, Associate, and Full Professor levels. Applications in all areas of accounting are invited. Applicants should send their *curriculum vitae*, three letters of recommendation, recent publications and working papers to: Ms. Karen Robertson, Faculty of Management, McGill University, 1001 Sherbrooke, St. West Montreal, Quebec, CANADA H3A 1G5. In accordance with Canadian Immigration requirements, priority consideration will be given to Canadian citizens and permanent residents of Canada. McGill University is committed to equity in employment.

307 Experienced PhD, CPA seeks a visiting position to teach financial, managerial, or international accounting courses for the 2006-2007 academic year. To this assignment, I bring a wealth of national and international teaching experience, an established publication record, and excellent teaching evaluations. Teaching assignments include faculty positions in the U.S. and the U.K. I participate regularly in AAA meetings and am a Regional Coordinator for the International Section. A *curriculum vitae* will be furnished upon request. You may contact me at profran@hotmail.com.

308 MISSISSIPPI STATE UNIVERSITY, College of Business and Industry, School of Accountancy, invites applications for an Assistant or Associate Professor of Accounting, a tenure-track position beginning Fall 2006. Responsibilities include teaching undergraduate and graduate courses in Taxation. Doctorate degree in accounting is required. The position requires excellence in teaching based on prior experience as a Graduate Teaching Assistant or as an Instructor or Assistant Professor, ability to do research leading to publication in nationally refereed journals, and an interest and willingness to be involved in professional accounting activities. An interest in advising and mentoring students is expected, including participation in the Ph.D. program. Professional certification and professional accounting experience is highly desirable. Salary is competitive. Send letter of application, current *curriculum vitae*, and names and contact information of three references to: School of Accountancy, Attn: Search Committee Chair, PO Box EF, Mississippi State, MS 39762. MSU is an AA/EOE.

309 COLLEGE OF MOUNT ST. JOSEPH invites applications for a full-time tenure track Assistant Professor in Accounting in the Department of Business Administration beginning Fall 2006. Candidates with a doctorate in Accounting are preferred; ABDs will be considered but must complete the degree prior to start date. Required competencies include demonstrated effectiveness in college teaching in Accounting Principles, Intermediate Accounting, and other courses in the applicant's area of interest; and demonstrated ability to successfully conduct scholarly research and publish, consistent with AACSB standards. Preference will be given to individuals with the ability to enhance teaching effectiveness with use of relevant technology. Applicants should send a cover letter, curriculum vita, evidence of teaching effectiveness, and arrange for three letters of recommendation to be sent to: Mr. Jim Brodzinski, Chair, Department of Business Administration, College of Mount St. Joseph, 5701 Delhi Rd., Cincinnati, OH 45233, or by email: Jim_Brodzinski@mail.msj.edu.

310 EAST CAROLINA UNIVERSITY, Department of Accounting, invites applications for a tenure-track Assistant/Associate Professor position in Accounting for the 2006-2007 academic year. Applicants should have or expect to have a Ph.D. in Accounting by August 2006, and appropriate professional certification. Our current needs include Government Accounting and Internal Auditing. A 12-s.h. load is standard with release time granted for research and publication productivity. Send credentials (cover letter which describes specific skills and experience relevant to this position, current resume/*vita*, evidence of research/publication activity, current teaching evaluations, and three original letters of reference) to Dr. Dan Schisler, Chair, East Carolina University, 3208-A Bate Building, Greenville, NC 27858-4353. Applicants must comply with the Immigration Reform and Control Act. ECU is an Equal Opportunity/Affirmative Action University and accommodates individuals with disabilities. Individuals requesting a disability accommodation should call the ECU Office Disability Support Services at (252) 328-6799 (Voice/TTY/Relay). Official transcripts are required upon employment.

311 TRINITY UNIVERSITY invites applications for the position of Jesse H. Jones Distinguished Professor of Accounting effective August 2007. The position requires a Ph.D./D.B.A. in Accounting, an established record of research, a commitment of continued scholarship, and excellence in teaching. Candidates should have the qualifications necessary for appointment as a Professor and will be expected to have an interest in teaching Financial Accounting Theory and Regulation. The teaching load is two courses per semester. The salary is competitive. The AACSB-accredited Department of Business Administration offers bachelor's and master's degrees in Accounting. Applicants should send a letter of interest, *curriculum vitae*, and a sample of written work to: Dr. Sankaran Venkateswar, Chair of Search Committee, Department of Business Administration, Trinity University, One Trinity Place, San Antonio, TX 78212-7200; Email: svenkate@trinity.edu. Review of applications will begin immediately, and the position will remain open until filled. Trinity University is an Equal Opportunity Employer.

312 UNIVERSITY OF NORTH CAROLINA, GREENSBORO, Bryan School of Business and Economics, Department of Accounting, invites applications for anticipated positions at the Assistant, Associate, or Professor levels, beginning August 1, 2006. Our primary needs are in financial, accounting information systems, and auditing, but all will be considered. A doctorate in accounting or an ABD nearing completion is required. Interested candidates should submit a letter of application, *vita*, and the names of three references to: Professor Bill Harden, Chair, Search Committee, PO Box 26165, Greensboro, NC 27402-6165. Review of applications will begin September 1, 2005 and continue until the position is filled. EEO/AA.

313 UNIVERSITY OF NORTH CAROLINA, GREENSBORO, Bryan School of Business and Economics, Department of Accounting, invites applications for anticipated positions at the Assistant, Associate, or Professor levels, beginning August 1, 2006. Our primary needs are in financial, accounting information systems, and auditing, but all will be considered. A doctorate in accounting or an ABD nearing completion is required. Interested candidates should submit a letter of application, *vita*, and the names of three references to: Professor Bill Harden, Chair, Search Committee, PO Box 26165, Greensboro, NC 27402-6165. Review of applications will begin September 1, 2005 and continue until the position is filled. EEO/AA.

314 THE BRATTLE GROUP provides consulting services and expert testimony on economic, finance, regulatory, and strategic issues. We are seeking an Associate to join our Finance consulting practice area in the Cambridge, Massachusetts office. The ideal candidate is a recent graduate with an Accounting Ph.D. or an Accounting Professor interested in a full-time consulting career. Candidates may be a CPA or have similar experience coupled with a recent M.B.A. degree. Practical experience should be in the areas of financial reporting and disclosure. To be considered, please visit the Careers section of our website and submit your materials there: http://www.brattle.com/Careers/Default.asp. Alternatively, materials may be faxed to our Washington, D.C. recruiting office at: (202) 419-3335. The Brattle Group is an Equal Opportunity Employer.

315 ILLINOIS WESLEYAN UNIVERSITY, Department of Business Administration, invites applications for two tenure-track positions in Accounting at the rank of Assistant Professor. The normal teaching load is three courses and two preparations per semester. Illinois Wesleyan seeks candidates with doctorates in accounting. ABDs nearing completion and others with extraordinary qualifications are also encouraged to apply. Starting date is August 2006. Salaries will be competitive. Illinois Wesleyan University has distinguished itself as an independent, liberal arts university. To apply, send a letter of application, *curriculum vitae*, three reference letters, and a transcript. In the application letter the successful candidate should be able to demonstrate a commitment to teaching business within a liberal arts setting. Review of applications will begin immediately. Send applications to: Robert A. Kearney, Chair, Department of Business Administration, Illinois Wesleyan University, PO Box 2900. Bloomington, IL 61702-2900.

316 LONG ISLAND UNIVERSITY, BROOKLYN CAMPUS invites applications for a full-time tenure-track faculty position in Accounting at the rank of Assistant Professor beginning Fall 2006 in the Department of Accounting, Taxation and Law. We seek an individual whose scholarly attainments are consistent with tenure-track expectations and who is committed to high-quality teaching, research, and professional service. Minimum Qualifications: Ph.D. in Accounting is required. Or, a Master's Degree in Accounting and a Ph.D. in a related field are required. Candidates with current CPA, CMA, or CIA are preferred. (ABDs will also be considered with an obligation to promptly complete all degree requirements). Teaching experience at the undergraduate and/or graduate levels is desired. The Application Process: Please attach letter of interest and *c.v.* to an email with subject heading "Tenure-track Faculty Position in Accounting" and send it to: Roberta.Scotto@liu.edu; or, send a hard copy to: Search Committee/Accounting Faculty Position, c/o Ms. Roberta Scotto, Department of Accounting, Taxation and Law, Long Island University, One University Plaza–H700, Brooklyn, NY 11201.

317 FLORIDA INTERNATIONAL UNIVERSITY, School of Accounting, is seeking faculty members for visiting professor positions for Fall 2006 and/or Spring 2007 to teach undergraduate- and graduate-level tax courses. Appointments at the Assistant, Associate, and Full professor ranks will be considered. Positions will be open until filled. Ph.D., D.B.A. or equivalent terminal degree required. A high degree of competence is expected in teaching at both the undergraduate and graduate levels. Professional certification is preferred. Salary is competitive and negotiable based upon qualifications and teaching load. The School is separately accredited by the AACSB. Undergraduate tax courses are taught at the University Park campus and Biscayne Bay campus. Graduate tax courses are taught on Saturdays at Biscayne Bay campus, and on Monday and Tuesday mornings at Downtown Miami campus. Contact: Cherie J. Hennig, Faculty Director, MST Program, School of Accounting, FIU, Miami, FL 33199; Phone: (305) 348-3266, Fax: (305) 348-2914, Email: hennigc@fiu.edu.

318 UNIVERSITY OF MINNESOTA, DULUTH, Labovitz School of Business and Economics, has a tenure-track position in health care management, beginning Fall 2006. For a complete position description and information on how to apply, please visit our website at http://www.d.umn.edu/sbe/faculty/openpositions.php. The University of Minnesota is an Equal Opportunity Educator and Employer.

319 UNIVERSITY OF ALABAMA, HUNTSVILLE invites applications for a tenure-track Assistant/Associate Professor of Accounting starting August 2006. Teaching specialization open but preference given to Managerial, Cost, Governmental, or Tax Accounting. Candidates must hold Ph.D. in Accounting from an AACSB-accredited institution. Associate Professor requires established record of research and quality teaching at both graduate and undergraduate levels. Previous professional experience is desirable. The business school is AACSB-accredited and has Compustat and Securities return dataset. Nestled in the foothills of the Appalachian Mountains, the population of the area is 350,000. The area has award-winning schools, excellent transportation, and numerous cultural and recreational opportunities. Send letter of interest, *vita*, and names and addresses of three references to: Dr. Dorla Evans, Chair of Accounting Search Committee, College of Administrative Science, UAH, Huntsville, Al 35899. Refer position number 06/07-182 on all correspondence. Review of applications begins immediately. The position will remain open until filled. UAH is an AA/EEO institution.

320 TEXAS TECH UNIVERSITY, Area of Accounting of the Rawls College of Business, invites applications for a one-year visiting position at the rank of Assistant or Associate Professor to begin in the fall semester of 2006. Our primary area of need is Financial Accounting: Principles of Financial Accounting and Intermediate Accounting. Preference will be given to candidates with an earned doctorate in accounting, but ABDs will be considered. Applicants should send a cover letter, *curriculum vita*, and a list of three references to: Dr. Ralph Viator, Search Committee Chair, Texas Tech University, Rawls College of Business, Lubbock, Texas 79409-2101. Alternately, materials may be sent electronically to: Ralph.Viator@ttu.edu.

321 WESTERN ILLINOIS UNIVERSITY, Department of Accountancy, invites applications for Chair of the Department. The chair provides leadership in the areas of teaching, research, and service; represents the interests of the Department to the College; and facilitates fundraising and maintenance of AACSB accreditation. The position is at the rank of tenured Full Professor, and excellent teaching, research, and service credentials are required. Salary is AACSB competitive. Appointment may commence July 1, 2006, January 1, 2007, or July 1, 2007. A letter of application, current *vita*, and the names and contact information of at least three references should be sent to: Accountancy Chair Search Committee, Dean's Office, College of Business and Technology, Western Illinois University, 1 University Circle, Macomb, IL 61455-1390. A full position description is available at http://www.wiu.edu/employment/emp.sphp?id=335.

322 MONTANA STATE UNIVERSITY, College of Business, seeks applicants to fill one tenure-track position in Accounting for Fall 2006. Montana State University has an enrollment of approximately 12,000 students and is located in Bozeman, an attractive, culturally alive small city, 90 miles north of Yellowstone National Park. Fully accredited by the AACSB, the College of Business offers a Bachelor of Science in Business with options in Accounting, Finance, Management, and Marketing, and a Master of Professional Accountancy. Colleagues in all areas of accounting will be considered, however, our greatest teaching needs are in the areas of accounting theory/research and systems. In addition, research activities and service consistent with the College's mission are expected. The position requires a Ph.D. or D.B.A. in Accounting or a closely related field from an AACSB-accredited university. Preference will be given to those who have completed their doctoral degrees or will complete them prior to beginning work at Montana State University. For more information and complete job descriptions with application requirements and deadlines, please visit: http://www.montana.edu/msuinfo/jobs/faculty/ or contact: Anne L. Christensen at annec@montana.edu or (406) 994-2043. ADA/AA/EO/Veteran's Preference.

323 UNIVERSITY OF COLORADO AT DENVER, Business School and Health Sciences Center, seeks applicants for two tenure-track faculty positions in accounting. Minimum qualifications for Assistant Professor include a Ph.D. in accounting and a commitment to excellence in research and teaching. Candidates for advanced rank (Associate or Full Professor) must demonstrate significant accomplishments in research, teaching and service. Specialization in managerial accounting, auditing and fraud examination, or accounting information systems is preferred, though all areas will be considered. Applicants who demonstrate proficiency using new accounting or information technologies in their teaching are especially encouraged to apply. Review of applications begins immediately, and will continue until the positions are filled. Both appointments begin Fall Term 2006, or as negotiated. Applicants for Visiting Positions are also encouraged to apply. Both positions are dependent on funding availability. The Carnegie Foundation has just classified UCDHSC as RU/VH: Research Universities (very high research activity). For full description, please see: http://www.ucdhsc.edu/admin/hr/facpos.htm#asstprof. The Business School website is: http://www.cudenver.edu/Academics/Colleges/Business/default.htm. Contact: Professor Bruce R. Neumann, Director of Accounting Programs, PO Box 173364, Campus Box 165, Denver, CO 80217-3364; or Bruce.Neumann@cudenver.edu.

324 UNIVERSITY OF LOUISIANA, MONROE invites applications for a tenure-track Assistant/Associate Professor of Accounting with a primary teaching emphasis in auditing, starting August 2006. The College of Business at ULM promotes a balance between teaching and intellectual contributions, combined with professional service to the university, discipline, and community. The Accounting Program and the College of Business Administration are both accredited by AACSB International. Candidates must have an earned doctorate (or ABD nearing completion) with appropriate professional certification and/or experience. Applicants should provide evidence of excellence in teaching and scholarship, *curriculum vitae*, transcripts, and names of three references. Applications will be considered until the position is filled; direct applications to Dr. Michelle McEachern, The University of Louisiana at Monroe, 700 University Avenue, Monroe, LA 71209-0110. The University is an Equal Opportunity/Affirmative Action Employer located in a metropolitan area of 150,000 persons. For more information, visit the ULM website: http://www.ulm.edu.

325 ALABAMA STATE UNIVERSITY, College of Business Administration, Department of Accounting and Finance, invites applications for an Assistant/Associate Professor of Accounting (JC-2022A). This is a nine-month, tenure-track position for the 2006 Fall Semester. The successful applicant will teach at the undergraduate and graduate level, engage in research and publication, participate in university-wide service and professional activities, mentor and academic advising of students. Qualifications: Ph.D. in Accounting from an AACSB International accredited school. Demonstrate research and publication in refereed journals. Strong leadership, administrative skills, and ability to work effectively in a shared governance environment is preferred. Previous faculty experience is required. Prior business experience and ability to interact with the business community are desirable. Salary: Competitive, Starting date: 2006 Fall Semester, immediately following selection and approval of a candidate. To apply: Interested applicants should submit an employment application form listing the names, address and phone numbers/email address of the last three employers on the reference sheet, transcripts, and resume to: Office of Personnel and Human Relations, Alabama State University, PO Box 271, Montgomery, AL 36101.

326 UNIVERSITY OF TEXAS–PAN AMERICAN invites applications for Department Chair of Accounting and Business Law to begin Fall 2006. Salary is AACSB competitive and commensurate with achievements and experience. Successful candidate should be eligible for tenure at rank of Associate or Full Professor. Candidates should have doctorate degree in accounting, strong leadership skills, ability to work and communicate effectively in a participative, collegial, culturally diverse environment, and willingness to develop and maintain fruitful relationships with professionals in the community. For more information please visit: http://www.utpa.edu/humanresources/employment/coba.html. Interested candidates should send a letter of interest and comprehensive *vita* to: Dr. Rubik Atamian, Search Committee Chair, Department of Accounting and Business Law, COBA, UTPA, 1201 W. University Drive, Edinburg, TX 78541-2999. Review of applications begins immediately, and position will remain open until filled. UTPA is an Equal Opportunity/Affirmative Action Employer. Women and minorities are encouraged to apply.

327 DeVRY UNIVERSITY seeks an Accounting/Finance Curriculum Manager and Full-time Faculty for Accounting and Finance courses. Individual would be responsible for both development and management of curriculum and faculty, plus teaching. The development component includes responsibility for course curriculum guides and web-related course development where "Best Practices" are utilized. Curriculum managers work with staff, faculty, and publishers to help assure that the curriculum is current with existing business practices and one of high quality. Management of curriculum includes approval of new faculty plus working with these individuals as they prepare for their first teaching experience for KGSM/DeVry University. Specific requirements: Earned doctorate plus academic and professional accounting and/or finance expertise; understanding of curriculum development processes; excellent teaching practices; computer literacy; excellent communication skills. Send resume and contact information for three references to: Dori Lombard at dlombard@keller.edu.

ST. MARY'S UNIVERSITY

Bill Greehey School of Business
St. Mary's University

ST. MARY'S UNIVERSITY, Bill Greehey School of Business invites applications to fill a tenure-track position in accounting at the assistant/associate/full professor level. Discipline expertise in auditing and information systems strongly desired. Expertise in other area(s) of accounting is a plus. Candidates are required to hold a PhD or DBA from an AACSB-accredited institution, and demonstrate the ability to bring excellence and innovation into the classroom, maintain a strong and continuing record of scholarship consistent with the teacher-scholar model, and possess attributes leading to successful professional service. Salary is competitive. The position begins as early as Fall 2006. Interested candidates should send an application letter, current vita, and names with contact information for three references to: Dr. Tom Madison, Chair, Department of Accounting, Bill Greehey School of Business, St. Mary's University, One Camino Santa Maria, San Antonio, TX 78228-8607. Deadline is June 1, 2006.